Language Contact

HSK 45.1

Handbücher zur Sprach- und Kommunikationswissenschaft

Handbooks of Linguistics and Communication Science

Manuels de linguistique et des sciences de communication

Co-founded by Gerold Ungeheuer
Co-edited by Hugo Steger 1985–2001

Founded and edited by
Herbert Ernst Wiegand 1982–2018

Band 45.1

De Gruyter Mouton

Language Contact

An International Handbook

Edited by

Jeroen Darquennes
Joseph C. Salmons
Wim Vandenbussche

De Gruyter Mouton

ISBN 978-3-11-044106-2
e-ISBN (PDF) 978-3-11-043535-1
e-PUB (EPUB) 978-3-11-043325-8
ISSN 1861-5090

Library of Congress Control Number: 2019935642

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

© 2019 Walter de Gruyter GmbH, Berlin/Boston
Cover design: Martin Zech, Bremen
Typesetting: Meta Systems Publishing & Printservices GmbH, Wustermark
Printing: CPI books GmbH, Leck

www.degruyter.com

Preface

This work has been a long time coming. All of us editors had the 1996/1997 HSK handbook on contact linguistics in our reference library within hand's reach and we must have used it countless times to touch base while writing, preparing lectures and classes, or looking for the ongoing and untamable inspiration and creativity of the handbook's editor-in-chief, Peter Nelde.

Ever since the turn of the decade, the suggestion of a reworked edition popped up at regular times. Jeroen was the logical 'volunteer' for the job, having taken over the torch when Peter, his *Doktorvater*, had to pass on the flame. A summer evening reception in De Gruyter's Berlin stomping grounds in the summer of 2012 did the final trick. The suggestion of a new handbook floated over talks in the inner courtyard, former Editorial Director Uri Tadmor said "You should do it", former Managing Director Anke Beck raised her glass and emphatically said "Yes". It took another year and a half to come up with a 'cunning plan' and to establish an editorial team.

Weighing the potential of the editorial job as one of the wilder and more sleep depriving rides in scholarship, Jeroen called on Joe and Wim among the rare breed of academic friends to form a complementary masthead for the project. The three of us fine-tuned the structure of the handbook in no time, soon agreed on a list of potential authors, started to invite contributors towards the end of 2015, and agreed on a 'modus operandi': given the absence of personal secretaries in today's academia, the editorial headquarters would be located in Jeroen's Macbook travelling back and forth between Tienen, Namur and the French Alps with fully functional and much more than just trustful lifelines attached to Joe's 24/7 Macbook in Wisconsin and Wim's wandering Macbook eventually getting a strong foothold in Asse. All the rest is a few years of history, perspiration, true editorial companionship and quality control, for which we don't want to hide our pride and determination.

Each article was read by all three editors, before going through double blind peer review by individually selected specialists. Authors received two reviews and a synthesis of those reviews completed with the editors' comments and with clear instructions for revision and/or rewriting. The revised versions were meticulously checked against the review summaries by the three editors (including conforming all bibliographies to the Handbook stylesheet), in view of submitting a third fully revised version with all remaining minor or major changes required. These submissions went through final correction and streamlining for spelling and style by a professional external editor, resulting (after a final check by the individual authors) in a coherent and polished submission for the production of galley proofs. As per usual, minor final corrections were administered in this final phase. The turnover time between first (beginning of 2017) and very final versions (very beginning of 2019) was about two years − shorter than usual for this type of volume, but well worth the investment.

None of this would have taken off without the patient and thorough input from the following reviewers, who swiftly helped us out on top of all other academic duties for no further compensation than a sincere word of thanks:

Philipp Sebastian Angermeyer	*York University*
Colin Baker	*Bangor University*
Peter Bakker	*Aarhus University*
Marlyse Baptista	*University of Michigan*
Elisabeth Barakos	*Aston University*
Alexander Bergs	*Universität Osnabrück*
Jelke Bethlehem	*Universiteit Leiden*
Robert Blackwood	*University of Liverpool*
Katharina Brizić	*Albert-Ludwigs-Universität Freiburg*
Joshua R. Brown	*University of Wisconsin − Eau Claire*
Rebecca Childs	*Coastal Carolina University*
Joseph Clancy Clements	*Indiana University*
Kathy Conklin	*The University of Nottingham*
Silvia Dal Negro	*Free University of Bozen-Bolzano*
Boyd Davis	*University of North Carolina − Charlotte*
Kees de Bot	*Rijksuniversiteit Groningen*
Jean-François de Pietro	*Université de Neuchâtel*
Alfonso del Percio	*University College London*
Vittorio Dell'Aquila	*Centre d'études linguistiques pour l'Europe, Milano*
Till Dembeck	*Université du Luxembourg*
Kati Dlaske	*University of Jyväskylä*
Robin Dodsworth	*North Carolina State University*
Stefan Dollinger	*The University of British Columbia*
Theodorus du Plessis	*Universiteit van die Vrystaat*
Lise Dubois	*Université de Moncton*
Robert Dunbar	*The University of Edinburgh*
Philip Durkin	*University of Oxford*
David Eddington	*Brigham Young University*
Andrea Ender	*Universität Salzburg*
Mel M. Engman	*University of Minnesota*
Geneviève J. Escure	*University of Minnesota*
Nicholas Evans	*Australian National University*
Floyd J. Fowler, Jr.	*University of Massachusetts − Boston*
Ben Frey	*University of North Carolina − Chapel Hill*
Susan Gal	*The University of Chicago*
Francesco Gardani	*Universität Zürich*
Paul B. Garrett	*Temple University*
Roberto Garvía	*Universidad Carlos III de Madrid*
Jessica Gasiorek	*University of Hawai'i at Mānoa*
Annette Gerstenberg	*Universität Potsdam*
Federico Gobbo	*Universiteit van Amsterdam*
Kellie Gonçalves	*University of Oslo*
Gabriel González Núñez	*The University of Texas − Rio Grande Valley*
Gilles Grenier	*Université d'Ottawa*
Jacques Guyot	*Université de Paris 8 − Vincennes*
Robert Hartsuiker	*Universiteit Gent*
Nancy Hawker	*University of Oxford*

Gabriele Hogan-Brun *University of Bristol*
Michael Hornsby *Adam Mickiewicz University, Poznań*
Francis Hult *Lund University*
Ellen Hurst *University of Cape Town*
Gabriele Iannàccaro *Università degli studi di Milano-Bicocca*
Mike Iverson *Indiana University*
Rudi Janssens *Vrije Universiteit Brussel*
Adam Jaworski *The University of Hong Kong*
David Cassels Johnson *The University of Iowa*
Barbara Johnstone *Carnegie Mellon University*
Elin Haff Gryffudd Jones *Aberystwyth University*
John E. Joseph *The University of Edinburgh*
Nina Kazanina *University of Bristol*
Remco Knooihuizen *Rijksuniversiteit Groningen*
Leena Kolehmainen *University of Turku*
Georg Kremnitz *Universität Wien*
Sol Lago *University of Potsdam*
Jarmo Lainio *Stockholms universitet*
Alfred Lameli *Albert-Ludwigs-Universität Freiburg*
M. Lynn Landweer *Dallas International University*
Epp Lauk *University of Jyväskylä*
Brian Lennon *Penn State University*
Erez Levon *Queen Mary University of London*
Paul Lewis *SIL International*
Christopher Lucas *SOAS University of London*
Friederike Lüpke *SOAS University of London*
Vally Lytra *Goldsmiths University of London*
Jeff MacSwan *University of Maryland*
Miki Makihara *The City University of New York*
Philip McDermott *Ulster University*
Annabelle Mooney *University of Roehampton*
Máiréad Moriarty *University of Limerick*
Silvia Moscatelli *Università di Bologna*
Frédéric Moulène *Université de Strasbourg*
Simon Musgrave *Monash University*
Nico Nassenstein *Johannes Gutenberg-Universität Mainz*
David A. Natvig *University of Wisconsin − Madison*
Nicole Nau *Adam Mickiewicz University, Poznań*
Luci Nussbaum *Universitat Autònoma de Barcelona*
Pádraig Ó Riagáin *Trinity College Dublin*
Leigh Oakes *Queen Mary University of London*
Lourdes Ortega *Georgetown University*
Jan-Ola Östman *University of Helsinki*
Andrew J. Pantos *Metropolitan State University of Denver*
John Paolillo *Indiana University*
Helma Pasch *Universität zu Köln*
Peter L. Patrick *University of Essex*

Yael Peled	*McGill University*
Falco Pfalzgraf	*Queen Mary University of London*
Nicolai Pharao	*Københavns Universitet*
Katja Ploog	*Université d'Orléans*
Sebastian Rasinger	*Anglia Ruskin University*
Sally Rice	*University of Alberta*
Gijsbert Rutten	*Universiteit Leiden*
Karène Sanchez	*Universiteit Leiden*
Thomas Schmidt	*Leibniz-Institut für Deutsche Sprache*
Inge Sichra	*Universidad Mayor de San Simón, Cochabamba*
Jeff Siegel	*University of New England*
Ellen Simon	*Universiteit Gent*
Marián Sloboda	*Univerzita Karlova, Praha*
Massimiliano Spotti	*Tilburg University*
James N. Stanford	*Dartmouth College*
Andrée Tabouret-Keller	*Université de Strasbourg*
Sara A. Trechter	*California State University − Chico*
Jeanine Treffers-Daller	*University of Reading*
Christian Uffmann	*Heinrich-Heine-Universität Düsseldorf*
Matthias Urban	*Universiteit Leiden*
Luk Van Mensel	*Université de Namur*
Mieke Vandenbroucke	*Universiteit Gent*
Martine Vanhove	*LLACAN, CNRS − Inalco*
Anna Verschik	*Tallinn University*
Cécile Vigouroux	*Simon Fraser University*
Julio Villa-García	*The University of Manchester*
Rik Vosters	*Vrije Universiteit Brussel*
Melanie Wagner	*Université du Luxembourg*
Keith Walters	*Portland State University*
Ingo Warnke	*Universität Bremen*
Jean-Jacques Weber	*Université du Luxembourg*
Iwar Werlen	*Universität Bern*
Søren Wichmann	*Universiteit Leiden*
Roland Willemyns	*Vrije Universiteit Brussel*
Colin H. Williams	*Cardiff University*
Margaret E. Winters	*Wayne State University*
José Woehrling	*Université de Montréal*
Jan Wohlgemuth	*Independent researcher*
Wolfgang Wölck	*University at Buffalo*
Heather Harris Wright	*East Carolina University*
Eline Zenner	*KU Leuven*
Klaus Zimmermann	*Universität Bremen*

We thank our contributors for generously accepting our suggestions, changes and endless requests for clarification, especially regarding our post-editing of spelling and punctuation for the sake of consistency. In the rare occasion where authors insisted on maintain-

ing their personal language norms, we hope the reader will appreciate the added value of authenticity.

De Gruyter's Berlin office remained our faithful fan base from the very onset of the project, sweetened with breakfasts or just some cups of coffee at Café Einstein or lunch at Valentino's on the corner of the Derfflingerstraße and the Kurfürstenstraße. A warm thanks to Birgit Sievert, Barbara Karlson and the rest of the editorial team for gently guiding us through the whole editorial process. Our proofreader, Brandy Trygstad, also deserves a special word of thanks. The speed, the patience and the accuracy with which she handled the chapters have been crucial help to us. We send gratitude and admiration to Helen Bilton, who compiled the index with incomparable skill and impressive speed. Jeroen also would like to thank (and pay tribute to) Wolf Wölck who was part of the editorial team of 1996 and whose friendship and subtle hints have had a much stronger impact on the handbook than he can ever imagine.

A final thought goes to our home bases who were our rock throughout the whole editorial journey. A warm, long and heartfelt embrace for our wives and families, who endured our absence and relentless distraction over the past years. And, yes, we promise that the second volume of this handbook will be the last of our mad projects − for some years to come.

Jeroen, Joe & Wim
February 2019

Contents

Preface v

1. Language contact research: scope, trends, and possible future directions · Jeroen Darquennes, Joe Salmons and Wim Vandenbussche 1

I. Linguistic aspects of language contact

2. Contact-induced change · Frank Seifart 13
3. Universal processes in contact-induced syntactic change · Sandro Sessarego 24
4. Contact-induced semantic change · Patience Epps and Danny Law 38
5. Lexicon and word formation · Stavroula Varella 52
6. Morphology · Felicity Meakins 63
7. Orthography and graphemics · Federica Guerini 76
8. Levels of representation in phonetic and phonological contact · David Natvig 88
9. Pidginization and creolization · Bettina Migge 100
10. Varieties in contact · James Wilson 112
11. Language contact and constructed languages · Marc van Oostendorp 124

II. Language contact and the individual

12. Pragmatics · Emilee Moore and Marco Santello 136
13. Borrowing · Yaron Matras 148
14. Code-switching · Gerald Stell 159
15. Language contact across the lifespan · Guro Nore Fløgstad and Elizabeth Lanza 172
16. Uninstructed language acquisition in multiple language learners · Annick De Houwer 185
17. First language attrition and contact linguistics · Gülsen Yılmaz and Monika S. Schmid 198
18. Individual variation in bilingual lexical processing: the impact of second language proficiency and executive function on cross-language activation · Janet G. van Hell, Katharine Donnelly Adams and Fatemeh Abdollahi 210
19. Metalinguistic awareness and multilingual development · Ulrike Jessner . . . 222
20. Language attitudes · Ian Bekker 234
21. Gender and language contact: how gender is/isn't marked in language contact · Miriam Meyerhoff and Alexandra Birchfield 246

III. Societal aspects of language contact

22. Multilingualism · Suzanne Romaine 257
23. From geographical and social boundaries to epistemic breaks · Robert Nicolaï 270
24. Language and identity in language contact settings · Luiza Newlin-Łukowicz 283
25. Language ideology · Kristine Horner and Andrew F. Bradley 296
26. Speech community · Julia Sallabank 307
27. Urban and rural language contact · Nkonko M. Kamwangamalu and Alla V. Tovares 319
28. Globalization and superdiversity · Lionel Wee 332
29. States, nations, and language contact · Alexander Maxwell 343
30. Language politics, policy, and planning · Heiko F. Marten 357
31. Majorities and minorities in language policy and language rights · Stephen May 370
32. Language standardization · Tore Kristiansen 384
33. Domains · Hartmut Haberland 397
34. Language contact and education: from language separation to translanguaging pedagogy · Anne-Marie de Mejía 406
35. Media and contact linguistics · Enrique Uribe-Jongbloed 418
36. Language contact in the linguistic landscape · Durk Gorter 431
37. Language shift · Maya Ravindranath Abtahian 441
38. Language maintenance · Matthias Brenzinger 454

IV. Methodological issues

39. Quantitative analysis of language contact data · Wladyslaw Cichocki 468
40. Qualitative data elicitation and analysis · Marie Maegaard, Karoline Kühl and Janus Spindler Møller 487
41. Surveys · Emilia Peytcheva 501
42. The representativeness of samples · Andreas Quatember 514
43. The geolinguistic treatment of demolinguistic data · Vittorio Dell'Aquila and Gabriele Iannàccaro 523
44. Domain analysis · Martin Ehala 536
45. Language typology in contemporary perspective · John Edwards 549
46. Social network analysis · Li Wei 561
47. Nexus analysis and multilingualism · Pia Lane 573
48. Ethnography in research on language contact · Teresa L. McCarty and Lu Liu 587
49. Interviewing · Michol F. Hoffman 600
50. Participant Observation · Ana Deumert 613
51. Discourse analysis · Mi-Cha Flubacher 625
52. Corpus linguistic methods · Evangelia Adamou 638
53. Psycholinguistic methods in the study of bilingualism · Peter A. Starreveld and Annette M. B. de Groot 653
54. Research ethics in contact linguistics · Keren Rice 668

V. Interactions with neighboring disciplines

55. Anthropology · Janet M. Fuller 682
56. Colonial studies · Doris Stolberg 695
57. Economics and language contact · François Grin 707
58. Educational sciences · Ofelia García 719
59. Language contact and legal studies · Eduardo J. Ruiz Vieytez 731
60. Contact linguistics and literary studies · David Gramling 743
61. Media/Communication studies · Tom Moring and Lia Markelin 754
62. The politics perspective on language contact · Peter A. Kraus 765
63. Social psychology · Diana Cárdenas, Laura French Bourgeois and Roxane de la Sablonnière 777
64. Translation studies · Dirk Delabastita 789

Index 803

1. Language contact research: scope, trends, and possible future directions

1. Introduction
2. Scope
3. Trends
4. Possible future directions
5. References

1. Introduction

Almost a quarter century ago, Mouton published a massive two-volume set called *Kontaktlinguistik / Contact Linguistics / Linguistique de Contact* (Goebl et al. 1996, 1997). The book you are reading now builds on that foundation, but even as we 'stand on the shoulders of giants', the field has progressed so rapidly and changed so dramatically that these two handbooks represent two fundamentally different books about very different fields. The present volume is not really a new edition of Volume 1 of the 1996 handbook in any sense, but instead a fresh collection on the same topic. This introduction provides a brief comparison of the two projects and situates the present volume in a broader context.

As pointed out by Oksaar (1996: 2), language contact phenomena have been "perceived and discussed in their various applied aspects throughout the history of Europe", and of course far beyond. The emergence of language contact as an area of scientific interest, however, basically dates back to the second half of the 19th and the first half of the 20th century, when Hesseling, Paul, Schuchardt, Whitney, and others developed an interest in dialect contact, pidgin and creole formation, and the roles of lexical, phonological, and syntactic transfer in language change (Clyne 1975, 1987; Winford 2007: 22). Also important to the later development of research on the linguistic consequences of contact between people and social groups was the work of Georg Schmidt-Rohr, Heinz Kloss, and other representatives of the *Auslandsdeutsche Volksforschung*, though Clyne (1987: 456) reminds us with reference to the work of Schmidt-Rohr that this work was "steeped in the racist ideology of National Socialism"; see also Hutton (1999). The bibliographies of Weinreich's ([1953] 1968) *Languages in Contact: Findings and Problems* and Haugen's ([1953] 1969) *The Norwegian Language in America* offer broad samples of early literature on intra- and extralinguistic language contact phenomena. During the time of the emergence of what Voegelin (s.d.) called 'hyphenated linguistics', both monographs are also generally considered to have sparked more systematic study of language contact phenomena (Matras 2009: 1). Particularly Weinreich's (1968) volume continues to inspire contemporary specialists, for example, Lim and Ansaldo (2016: 187–190).

Less discussed in recent literature, but an equally important source of inspiration for the present volume, is Mackey's (1976) *Bilinguisme et Contact des Langues*. The work of Mackey and his colleagues at the University of Laval's International Center for Research on Bilingualism inspired Nelde's (1980, 1983) take on language contact. As a consequence, it also had an influence on the scope and the structure of the 12th volume in De Gruyter's HSK-Series mentioned just above: the international and trilingual (English-

https://doi.org/10.1515/9783110435351-001

French-German) handbook on *Contact Linguistics*, edited by Peter H. Nelde (who acted as lead editor) together with Hans Goebl, Zdeněk Starý, and Wolfgang Wölck and published in 1996 (Volume 1) and 1997 (Volume 2). Thomason (2001: 13) refers to the handbook as a "major […] encyclopedic treatment of the subject". Rindler Schjerve (1999: 8) sees it as offering "an ultimate synthesis of contact linguistic endeavours". While this might be an overstatement, it cannot be denied that the thematic coverage of the field offered in the first, theoretical volume of *Contact Linguistics* is vastly broader than that offered in other, more recently published monographs and handbooks on language contact. At the time of this writing, the field has a rapidly growing set of handbooks beyond the original Mouton project, including Hickey (2013, new edition in preparation), Grant (forthcoming), and Mufwene and Escobar (in preparation), as well as textbooks and journals.

As will become clear in Section 2 of this introduction, the structure of and philosophy behind the 1996 volume have inspired the present volume in many ways. Section 3 is devoted to some current trends in language contact research. Section 4 closes with potential directions for future research.

2. Scope

Matras (2009: 1) lists borrowing, areal language clusters, new 'contact' languages, pidgins and creoles, and code-switching as topics that receive especially extensive attention in recent publications on language contact. Those are, for example, topics that are dealt with frequently in the *Journal of Language Contact* (JLC), a thriving platform for language contact research established by Robert Nicolaï in 2007. As stated on the journal's website, *JLC* particularly aims at "advancing our understanding of the nature of language" (https://brill.com/view/journals/jlc/jlc-overview.xml), which almost automatically entails a focus on the structural linguistic consequences of language contact rather than on speakers and social groups as the loci of language contact. In the opening article of the first issue, Nicolaï (2007: 17) does, however, point to the variety of intersecting angles from which language contact phenomena are or can be studied. In doing so, he refers to Winford's (2007: 22) view that "[t]he earliest conceptions of the field of Contact Linguistics envisioned it as a multi-disciplinary area of study, encompassing a broad range of language contact phenomena and issues, linguistic, sociolinguistic, sociological and psycholinguistic". That is indeed what emerged early on in the preface of *The Norwegian Language in America*, in which Haugen (1969: xi) writes: "It has been my intention throughout this book to treat language as a social phenomenon, in such a way that it might offer something of value to students of history and sociology". This is also the vision that Weinreich (1968: 4) presents in the first chapter of *Languages in Contact*:

> Purely linguistic studies of languages in contact must be coordinated with extra-linguistic studies on bilingualism and related phenomena. Geographers and ethnographers have described bilingual populations; sociologists have examined the functioning of coexisting languages in a community; jurists have studied the legal status accorded to minority languages in various states; the inquiries of educators interested in bilingual children and in foreign-language teaching have stimulated psychologists to analyze the effects of bilingualism on personality. All the studies are described in a vast, scattered literature. But divergent as

> they are in purpose and scope, they are all essentially complementary in understanding a phenomenon of so many dimensions.

Work on language change, and linguistic theory more broadly, has long wrestled with how to put material into discrete boxes of 'internal' versus 'external' motivations, i.e. structural versus social. These passages show that work on language contact has engaged with the problem in a more holistic way, increasingly avoiding what Dorian (1993: 152) calls "the weakness of simplistic dichotomous thinking". In this regard, other subfields of linguistics are finally catching up with language contact studies.

The broad, inclusive take on language contact research just sketched is reflected in the many working papers and publications of the International Center for Research on Bilingualism (1967–2001; known from 1990 onward as the Centre international de Recherche sur l'Aménagement linguistique), as well as in the 30 volumes of the *Plurilingua* series published by the Brussels-based Research Center on Multilingualism (1977–2007). The content of these publications can be classified according to the interrelated main foci of research on language contact identified by Clyne (1975, 1996), Mackey (1976), and Nelde (1992), namely: language, individual language user(s), and society. In line with this view, the first three sections of the present volume focus on these three areas, while the final two turn to issues of methodology and connections to neighboring disciplines.

> **Section 1** focuses on linguistic aspects of language contact, and contains ten chapters covering contact-induced change from a general point of view (Chapter 2), contact-induced syntactic change (Chapter 3), contact-induced semantic change (Chapter 4), lexicon and word formation (Chapter 5), morphology (Chapter 6), orthography and graphemics (Chapter 7), levels of representation in phonetic and phonological contact (Chapter 8), pidginization and creolization (Chapter 9), varieties in contact (Chapter 10), and constructed languages (Chapter 11).
>
> **Section 2** deals with language contact and the individual. Its ten chapters focus on pragmatic aspects of language contact (Chapter 12), borrowing (Chapter 13), code-switching (Chapter 14), age groups (Chapter 15), uninstructed language acquisition in multiple language learners (Chapter 16), first language attrition (Chapter 17), individual variation in bilingual lexical processing (Chapter 18), metalinguistic awareness and multilingual development (Chapter 19), language attitudes (Chapter 20), and gender (Chapter 21).
>
> **Section 3** covers societal aspects of language contact. It starts with chapters on multilingualism (Chapter 22), geographical and social boundaries (Chapter 23), language and identity (Chapter 24), and language ideology (Chapter 25). It further contains chapters on the notion of speech community (Chapter 26), on urban and rural language contact (Chapter 27), globalization and superdiversity (Chapter 28), states, nations, and language contact (Chapter 29), language politics, policy, and planning (Chapter 30), majorities and minorities in language policy and language rights (Chapter 31), and language standardization (Chapter 32). It ends with chapters on domains of language use (Chapter 33), education (Chapter 34), media (Chapter 35), linguistic landscapes (Chapter 36), language shift (Chapter 37), and language maintenance (Chapter 38).

Organizing these chapters into these three sections with a focus on language, individual language user(s), and society seemed more straightforward than trying to force the chapters into the structure of 1996's Volume 1, which contained sections on: levels of linguis-

tic structure (Section III; 8 chapters), external research (Section IV; 6 chapters), basic approaches to contact linguistics (Section V; 17 chapters), central issues in contact linguistics (Section VI; 45 chapters), and applied contact linguistics (Section VIII; 15 chapters).

The earlier Volume 1 also contains a section on empirical methods and procedures (Section VII; 13 chapters), on which we modeled the empirical section of the present volume. Yet we have chosen to include more specific (rather than more general) methodological chapters.

> **Section 4**, on methodological issues, contains two broad introductory chapters on quantitative (Chapter 39) and qualitative analysis (Chapter 40) of language contact data. It also contains two overview chapters written by social scientists on surveys (Chapter 41) and the representativeness of samples (Chapter 42), which also serve as background for a chapter offering a broad introduction to the geolinguistic analysis of demolinguistic data (Chapter 43). Furthermore, this section contains chapters on domain analysis (Chapter 44), ways to establish a sociology-of-language framework to compare minority languages (Chapter 45), social network analysis (Chapter 46), nexus analysis (Chapter 47), ethnography (Chapter 48), interviewing (Chapter 49), participant observation (Chapter 50), discourse analysis (Chapter 51), corpus linguistic methods (Chapter 52), psycholinguistic methods (Chapter 53), and research ethics (Chapter 54).

Inspired by Section II (9 chapters) of the 1996 handbook, the present volume also contains a number of chapters that offer a more 'external linguistic take' on the study of language contact phenomena.

> **Section 5** treats robust interactions with neighboring disciplines from the perspective of anthropology (Chapter 55), colonial studies (Chapter 56), economics (Chapter 57), educational sciences (Chapter 58), legal studies (Chapter 59), literary studies (Chapter 60), media and communication studies (Chapter 61), political science (Chapter 62), social psychology (Chapter 63), and translation studies (Chapter 64).

The chapters have been ordered simply by our sense of what might work for someone reading cover-to-cover, in the unlikely event that someone would actually do that. The exception is Section 5, which, as an overview of neighboring disciplines, was ordered alphabetically by field of study. Since all chapters are encyclopedic, stand-alone chapters, they can be read separately, and it will make sense to read them in various orders depending on the goals and background of the reader. Thematically related chapters in Sections 1–3 do create potentially useful clusters of chapters. Section 1, for example, contains a number of related chapters on structural aspects of contact-induced change (Chapters 2, 3, and 4). We can also imagine value in a 'combined reading' of the chapters on urban and rural language contact (Chapter 27), globalization and superdiversity (Chapter 28), and states, nations, and language contact (Chapter 29).

While we have tried to offer our readers a volume which is both condensed and comprehensive, with theoretical chapters on a great variety of language contact phenomena, there are gaps. Section 1 was supposed to contain a chapter on sign languages, and Section 2 chapters on language contact and the brain, intercultural communication, and multiliteracies. For Section 3, we had planned a chapter on migration, *Abstand* and *Ausbau*, and language conflict (given the inextricable relationship between language

contact and language conflict as expressed in Nelde's Law: there is no contact without conflict, Nelde 1987). Section 4 should have included a chapter on neurolinguistic methods, and Section 5 a chapter on history and a chapter on sociology. These gaps result from the usual academic and nonacademic 'facts of life', ranging from administrative duties that made it impossible to deliver on time, unforeseen dramatic personal circumstances, and our decision to give priority to the timely publication of the project rather than allowing a long postponement in the name of creating a more complete volume. While we will weave these missing topics into the second volume of this handbook, there is, of course, enough literature out there today that readers can consult to see the state of the art on these matters. The 1996 volume on *Contact Linguistics* provides an easy point of departure, one that has stood the test of time. Yet it is also clear that new trends have come into being, a topic to which we now turn.

3. Trends

As editors, we happily refrain from the common practice of summarizing individual chapter content to introduce a volume like this, but instead take the opportunity to offer a 'snapshot in action' of ongoing changes in our field of inquiry, to assess its vitality and agility, and to testify to its methodological innovations, engagement with societal change at large, and manifold interactions with neighboring disciplines.

To this end, we highlight a number of trends that consistently manifested themselves across chapters in the major sections, both as marked progress from the previous HSK handbook on contact linguistics and as harbingers for key future research directions. Above all, we underscore how profoundly different the field is today as compared to only a few years ago on every relevant parameter, from new amounts and kinds of data, to new theories about language structure and social dimensions of language, to methods and topics of research.

3.1. Linguistic aspects of language contact

The past two decades have witnessed a major leap in the study of structural aspects of language contact. Our understanding of contact-induced change has substantially deepened, regarding both contact in the distant past and change now in progress, and especially of the latter, driven by the rise of language variation and change as a field and inspired by Weinreich, Labov, and Herzog (1968). Synchronic and diachronic approaches have enhanced one another. The massive increase in available evidence, as well as methodological innovations allowing for more refined analyses of these data, have led to increased insights into the (ir)regularity of language change, allowing (among many other things) for more refined models of contact-induced change, far beyond the 1996 canon. Growing attention to the distinct processes of 'matter' and 'pattern' borrowing is but one illustration of this epistemological progress.

The basic theoretical underpinnings of grammar (be it phonology, morphology, or syntax) are increasingly tightly woven into the fabric of language contact theory proper, offering insights into grammatical contact effects that can be directly applied to both the

conversational role and the language proficiency of the individual speaker in a given interaction. The landscape in grammatical theory has itself been transformed since the original volumes appeared; while the Contrastive Hierarchy in phonology and Minimalism in syntax were only emerging 20–25 years ago, they are central to some discussions in this volume. The systematic study of semantic change (aside from grammaticalization and related research) barely existed in some sense, but has boomed recently, opening new avenues in the understanding of the 'mechanics' of linguistic systemic convergence and of the societal and cognitive practices involved.

All of these evolving insights into the structural properties of language contact in the strictest grammatical sense have in particular informed (and apparently also created new momentum in) the study of contact varieties and contact languages. This obviously holds true for the processes of pidginization and creolization, but also for other emerging 'mixed' varieties. Universal second-language acquisition strategies are now instrumentalized, for example, to account for current developments in creoles, complementing traditional work on (de)creolization or substrate effects. We note a major shift from purely systemic to more inclusion of sociohistorical and usage-based criteria. The integration of sociocultural (and other extralinguistic) practices in the study of writing systems for these languages exemplifies this evolution, and provides telling evidence of the growing integration of purely linguistic dimensions of language contact studies with the role of the individual speaker in the process.

3.2. Language contact and the individual

The chapters in this volume make many clear and convincing cases for a shift in the study of the role of the individual language user in language contact situations, both complementing and rivaling the impressive progress made regarding the systemic aspects proper of language contact. Compared to 1996, the notion of the 'multilingual repertoire' now dominates the multidisciplinary assessment of language use in interaction. Ethnographic and anthropological perspectives on 21st century communication seem to have inspired a widespread, and well-received, urge for terminological innovation when describing and categorizing the individual's variety choices in distinct settings.

The present volume bears witness to this new mainstream discourse with ample reference to an elaborate corpus of vogue terms, including *superdiversity*, *polylanguaging*, *translanguaging*, or even plain *languaging*. The chapters where these concepts are used to both reappraise and reframe dimensions of linguistic contact together present a concise overview of the potential longevity and impact of this rejuvenated jargon, well beyond the ideology-laden contexts where it appears as a fixed asset.

While the intricacies and added value of these new labels lead the initiated to a deeper understanding of previously unfathomed dimensions of contact, we as editors were nevertheless struck by the sheer amount of effort and explanation required to enlighten the linguistic community at large about the relevant differences separating two terminological generations. Looking towards the potential third incarnation of this handbook in a decade or two from now, we marvel at the outcome of this terminological inflation, as more eloquently and elaborately explicated by Pavlenko (2018), Grin (2018), and Meyerhoff (2019: 136–137), to name but a few.

A river of present-day scholarship foregrounds ideologies, attitudes, and agency – whether linguistic or not – that control and shape shifts in speakers' repertoires, turning conversation in contact situations into a sequence of enactment and performance, with the individual speaker as a continuously stimulated actor responding to highly context-specific impulses.

While the appeal of the new in some cases tends to eclipse (or depreciate) terms from the traditional sociolinguistic canon, like 'code-switching' and 'transfer', this transparent classic terminology also saw continued refinement, appealing to new developments in the neighboring disciplines of sociology and psychology.

As opposed to this terminological modernism, the study of language acquisition proper saw substantial advancement through the full integration of psycholinguistic insights and experimental methods. The emergence of longitudinal studies over the lifespan, the further exploration of instructed versus uninstructed acquisition, and the study of cross-linguistic factors in language loss and attrition are just a few of the many innovations that led to deeper understanding of language learning in contact situations. These insights now also inform new approaches to the study of systemic language change at large, and at the same time, re-dynamize scholarship on the societal aspects of language contact.

3.3. Societal aspects of language contact

It comes as no surprise that the reach and impact of the 'language ideological turn' transcended the domain of personal multilingualism, and percolated steadily into the study of societal aspects of language contact over the past two-plus decades. The super-diversity paradigm informs many new contributions to this field, indeed, even if it sometimes appears to border on infatuation. Classic methodological categories and terminology are in constant flux, as the locus of research shifts from languages and their communities of users to the ways in which these are embodied, created, and manipulated. Identity creation, stance-taking, and other forms of speaker agency are now at the heart of societal contact studies, in a clear shift away from more traditional concepts of distinct speech communities with clearly defined properties. The multilayered, quicksilver linguistic identity of the 21st century speaker is mirrored in a conceptualization of language communities which continuously adapt and change according to a multitude of social stimuli. Increased migration and globalization are most certainly the first among equals in this respect: whilst to the historically informed observer, current migration patterns may not be as novel as they are sometimes presented as being, this mobility is definitely the best attested and easiest to trace ever, including also in terms of its linguistic effects. We may, in other words, be in the most privileged position ever as linguists to observe, analyze, and contextualize processes of change 'in action' during the contact process. Whether approached from an ethnographic or anthropological perspective, relatively straightforward models of language communities gave way to the concept of a 'pixelated' society, with individuals that permanently redefine their (linguistic) position and nature. Bridging this collection of micro-studies of hyper-individual variation to a solid and relevant assessment of macro-level change remains a challenge. Interest in stance-taking and sociopolitical agency among scholars in this new playing field enhances the

position of the observing and analyzing linguist with a 'critical' guise, and certainly adds an extra layer to the defeat.

An equally influential transformation touches upon the very nature of language contact itself. The tsunami of 'new' internet-based media is changing our conception of communication, to the extent that language contact no longer requires the co-occurring physical presence of speakers. As such, the nature of processes of societal language attrition and loss, traditionally established in a context of actual real-life encounter and/or withdrawal, has been reconceptualized. It would be an oversimplification to state that speakers nowadays are in constant contact due to the inescapable intrusiveness of the internet beyond their own will; yet technological advancements may have had their biggest impact, so far, when it comes to facilitating and simplifying worldwide contact between language users.

On the other hand, given that only roughly half of the world's population has internet access, ethical considerations on a divide between the 'haves' and 'have-nots' can complement the observation that the more traditional approaches to language contact may still be applicable, and should not be discarded for some time to come.

3.4. Methodological issues

The great strides in linguistic methods generally are obvious enough, especially in the rapidly increasing quantitative and qualitative sophistication of the whole field, including in language contact studies. Those methods are now more varied and sophisticated than ever, and are applied across the whole gamut of available data, whether spoken, written, or signed, in both experimental and real-life settings.

The old dichotomy between quantitative and qualitative perspectives has given way to a set of combined-methods approaches, incorporating the best of both worlds to better understand how language contact works, and getting rid of the clichéd opposition between 'subjective' micro-studies and 'objective' figure-based macro-surveys. As a consequence, quantitative techniques themselves are now also a subject of qualitative consideration, including the ideological caveats inherent to specific methodologies. Accordingly, a persistent concern for the ethical dimensions of methodological choices has now become an acquired feature of research design. The clustering of analytical frameworks from various neighboring disciplines is one creative solution used to navigate possible pitfalls in this respect.

There is no doubt either that the material we work with has gotten substantially better as well. If the sheer amount of additional data available compared to the mid-1990s changed the scene, the organized ways in which this data is now available to researchers give an unprecedented boost to the field. To mention but one obvious example, corpus data and other large datasets make statistical analyses possible that could not be done with the far smaller amounts of data many early studies were built upon. Technological innovation was the driving force both enabling the compilation and facilitating the increasingly powerful manipulation of these 'big data', from high-level supercomputing to day-to-day routine data processing on commonly available and omnipresent devices.

These methodological and technical revolutions have had salient effects for all to see in a wide range of applications. The fine-grained mapping and monitoring of linguistic

diversity not only gathered momentum (and especially so in sensitive or traditional areas of language contact), but also had decisive impact on political and societal decisions relating to the communities concerned. Advances in imaging techniques have revolutionized the study of language processing, with impressive applications in healthcare and clinical research, but with equally groundbreaking progress in the psycho- and neurolinguistics of language contact.

These new horizons in subject matter (data) and subject method (analytical techniques) will shape the field in the years to come. We have included chapters and discussions on these, but also on traditional, concerns, stretching back to domain analysis. As noted above, the input from non-linguists in these discussions is much valued and vitally important, underscoring our methodological connections to allied disciplines.

3.5. Interactions with neighboring disciplines

While language contact studies are defined by interaction, exchange of information, and an integration in society at large, the discipline itself can only realize these properties in dialogue with neighboring fields of inquiry. No matter how fascinating the massive changes may have been over the past 25 years in the splendid isolation of our profession, the reality check for any of these should also be assessed by its reception across the humanities and social sciences, and increasingly across the rest of the scientific enterprise, such as biology and mathematics.

The role of languages-in-contact as a key factor of social identity has been overtly embraced by adjacent disciplines, including the speaker's autonomous capacity to reposition and redefine this very identity through linguistic variation. The multilingual and multimodal individual has by default become a core unit of analysis. The accompanying discourse on translanguaging may not (yet) pervade all of these proximate subject areas, but intergroup dynamics are − at least partially − increasingly analyzed along with individual linguistic profiling of the group members.

This observation extends to the scholar, in what seems to be the definitive embodiment of the 'linguistic' turn in the humanities over the previous century. While the fact that the medium of language colors the nature of scholarly opinions and views has been common core for decades, recent scholarship now redefines the complex relationships between researchers and their subject matters from the perspective of their meta-identity as active language users who are prone to all of the 'languaging' strategies discussed above.

Many other dimensions of the (socio)cognitive paradigm that have reinvigorated our discipline are now being equally applied in contact-related studies outside of the strictly linguistic sphere. A focus on language and diversity, and the possible societal and political consequences of linguistic inequality and/or injustice, now also informs much of the sociology- and education-related research in the social sciences writ large. A discourse of empowerment takes center stage in these disciplines, too, one which refers to language contact effects and is similar to the agency-driven positioning in contact linguistics discussed earlier.

The exponential spread of language contact through the 'digital revolution' is touched upon in various other domains. Considerations of both the strengthening and weakening effects of technology on minority groups, the spread of (language) ideologies through

technological innovation, and the default state of 24/7 connectedness have sparked scientific ambitions and concerns that both mirror those voiced by the contributors to this volume, and also guarantee an interdisciplinary interest in ongoing collaboration with contact linguists.

Overall, we see a pressing shared concern for the cultivation and preservation of a diverse language ecology worldwide, including the tension between maximizing personal expression, resisting commodification, and policing the fragile equilibrium between dominant and marginalized voices.

4. Possible future directions

If the next decades bring changes to our understanding of language contact of the magnitude we have witnessed since the original handbook, all bets are off about where the field will be. Predicting the precise paths of progress in distinct domains and on specific topics is tricky for any scholar, but it feels safe to say that the major trends affecting the daily lives of our speakers − be they political, societal, cultural, natural, or technological − will have an impact on language contact as well. As the second decade of the 21st century draws to a close, migration worldwide is on the rise for a variety of reasons − triggering new contact situations alongside political reactions aimed at both inclusion and isolationism. Shifting geo-economic dominance may complement and reinforce these changes with unprecedented vigor.

Advances in technology will likely push the limits of our analytical capacities well beyond the current state of affairs, be it in the realm of brain studies, cognitive insights, or data processing. As interdisciplinarity becomes the norm, part of the core of language contact research may actually be taken out of linguists' hands. The expected quantum leap in communication technology directed at mass use by the public at large is bound to redefine the very concept of language contact, especially when high-level scholarship and day-to-day-life intertwine in a combination that is anything but science fiction. More dramatically, facts such as climate change will create unprecedented contact situations. Linguists working on these issues may have even moral and ethical roles to play as the anticipated 'ecolinguistic crisis' sets in, with hundreds of now-endangered languages potentially wiped away in coming decades. More broadly, many changes are taking place in society that have and will continue to have impact on language contact as a phenomenon that colors the daily life in many increasingly hybrid speech communities around the globe. We hope that researchers − on the example of Fishman (1991) and others − will contribute to theoretical and empirical foundations to assist speech communities whose languages and whose very existence are threatened.

Judging from the insights and expertise gathered in this volume, we are confident that all concerned are ready to tackle these challenges, and that the shock of the new will not eclipse the knowledge and impact of insights developed in the earlier days of language contact studies proper. In that spirit, we sincerely hope that this handbook may serve as a gateway to language contact research, especially for a younger generation.

5. References

Clyne, Michael
1975 *Forschungsbericht Sprachkontakt*. Kronberg: Scriptor.

Clyne, Michael
1987 History of research on language contact. In Ulrich Ammon, Norbert Dittmar & Klaus J. Mattheier (eds.), *Soziolinguistik / Sociolinguistics*, vol. 1, 452–459. Berlin: De Gruyter.

Clyne, Michael
1996 Sprache, Sprachbenutzer und Sprachbereich. In Hans Goebl, Peter H. Nelde, Zdeněk Starý & Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact linguistics / Linguistique de contact*, vol. 1, 12–22. Berlin & New York: Walter de Gruyter.

Dorian, Nancy
1993 Internally and externally motivated change in language contact settings: Doubts about the dichotomy. In Charles Jones (ed.), *Historical linguistics: Problems and perspectives*, 131–155. London: Longman.

Fishman, Joshua A.
1991 *Reversing Language Shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon: Multilingual Matters.

Goebl, Hans, Peter H. Nelde, Zdeněk Starý & Wolfgang Wölck (eds.)
1996 *Kontaktlinguistik / Contact linguistics / Linguistique de contact*, vol. 1. Berlin & New York: Walter de Gruyter.

Goebl, Hans, Peter H. Nelde, Zdeněk Starý & Wolfgang Wölck (eds.)
1997 *Kontaktlinguistik / Contact linguistics / Linguistique de contact*, vol. 2. Berlin & New York: Walter de Gruyter.

Grant, Anthony P. (ed.)
Forthcoming *The Oxford handbook of language contact*. Oxford: Oxford University Press.

Grin, François
2018 On some fashionable terms in multilingualism research: Critical assessment and implications for language policy. In Peter A. Kraus & François Grin (eds.), *The politics of multilingualism: Europeanisation, globalisation and linguistic governance*, 247–274. Amsterdam: John Benjamins.

Haugen, Einar
1969 [1953] *The Norwegian language in America: A study in bilingual behavior*. Bloomington: Indiana University Press.

Hickey, Raymond (ed.)
2013 *The handbook of language contact*. Oxford: Wiley-Blackwell. [Second edition in preparation].

Hutton, Christopher M.
1999 *Linguistics and the Third Reich: Mother-tongue fascism, race and the science of language*. London: Routledge.

Lim, Lisa & Umberto Ansaldo
2016 *Languages in contact*. Cambridge: Cambridge University Press.

Mackey, William F.
1976 *Bilinguisme et contact des langues*. Paris: Klincksieck.

Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.

Meyerhoff, Miriam
2019 *Introducing sociolinguistics*, 3rd edn. New York: Routledge.

Mufwene, Salikoko S. & Anna María Escobar (eds.)
In preparation *The Cambridge handbook of language contact*. Cambridge: Cambridge University Press.

Nelde, Peter H.
1980 Einleitung. In Peter H. Nelde (ed.), *Sprachkontakt und Sprachkonflikt / Languages in contact and conflict / Langues en contact et en conflit / Taalcontact en taalconflict*, 1–2. Wiesbaden: Steiner.
Nelde, Peter H.
1983 Plaidoyer pour une linguistique des langues en contact. In Peter H. Nelde (ed.), *Gegenwärtige Tendenzen der Kontaktlinguistik / Current trends in contact linguistics / Tendances actuelles de la linguistique de contact* (Plurilingua I), 5–12. Bonn: Dümmler.
Nelde, Peter H.
1987 Language contact means language conflict. *Journal of Multilingual & Multicultural Development* 8. 33–42.
Nelde, Peter H.
1992 Multilingualism and contact linguistics. In Martin Pütz (ed.), *Thirty years of linguistic evolution*, 379–397. Amsterdam: John Benjamins.
Nicolaï, Robert
2007 Language contact: A blind spot in 'things linguistic'. *Journal of Language Contact* 1(1). 11–21.
Oksaar, Els
1996 The history of contact linguistics as a discipline. In Hans Goebl, Peter H. Nelde, Zdeněk Starý & Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact linguistics / Linguistique de contact*, vol. 1, 1–12. Berlin & New York: Walter de Gruyter.
Pavlenko, Aneta
2018 Superdiversity and why it isn't: Reflections on terminological innovation and academic branding. In Barbara Schmenk, Stephan Breidbach & Lutz Küster (eds.), *Sloganization in language education discourse: Conceptual thinking in the age of academic marketization*, 142–168. Bristol: Multilingual Matters.
Rindler-Schjerve, Rosita
1999 There is no contact without conflict. In Peter J. Weber (ed.), *Contact + confli(c)t: Language planning and minorities* (Plurilingua XXI), 1–12. Bonn: Dümmler.
Thomason, Sarah G.
2001 *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
Voegelin, Charles F.
s.d. On the probability of autonomous linguistics converging with hyphenated linguistics. Philadelphia, PA. [Unpublished manuscript].
Weinreich, Uriel
1968 [1953] *Languages in contact: Findings and problems*. The Hague: Mouton.
Weinreich, Uriel, William Labov & Marvin I. Herzog
1968 Empirical foundations for a theory of language change. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Directions for historical linguistics: A symposium*, 97–195. Austin: University of Texas Press.
Winford, Donald
2007 *An introduction to contact linguistics*. Oxford: Blackwell.

Jeroen Darquennes, Namur (Belgium)
Joe Salmons, Madison, WI (USA)
Wim Vandenbussche, Brussels (Belgium)

I. Linguistic aspects of language contact

2. Contact-induced change

1. Introduction
2. Prehistoric contact-induced change and linguistic universals
3. Asymmetries in contact-induced change
4. Modelling contact-induced change
5. Conclusions and outlook
6. References

1. Introduction

Contact-induced change is manifold in its manifestations and ubiquitous in the world's languages, suggesting that most humans were probably multilingual throughout most of human history. Understanding contact-induced change is thus crucial for explaining the historical development of language and the synchronically observable diversity of languages. A growing amount of empirical evidence for contact-induced change is available from diverse languages, and a range of models have been proposed to establish the limits on the range of the attested variability. Some of these models posit absolute constraints on contact-induced change, based on theoretical considerations and more or less restricted samples of languages. Other approaches start out by describing which kinds of changes are more probable in larger samples of languages in order to establish the relative strength and interaction of various factors involved.

In this chapter, I use traditional terminology such as 'recipient language', 'donor language', and 'borrowing'. Alternative terms for the latter are 'transfer' (Wilkins 1996; Myers-Scotton 2002) 'adoption' (van Coetsem 2000; Winford 2005), and 'copying' (Johanson 2008). The last is particularly attractive for two reasons: Firstly, it highlights that contact-induced changes are the result of creative processes of imitation. Secondly, the associated code-copying framework explicitly acknowledges that these processes involve a number of, to some extent independent, components, namely M(aterial), S(emantic), C(ombinatorial), and F(requential) aspects of the linguistic sign. The absence or presence of morphological material in borrowing is also captured in the terms MAT(ter) vs. PAT(tern) borrowing (Matras and Sakel 2007). Finally, contact-induced changes of a particular kind, namely those typically introduced by second-language learners of the recipient language (see Section 4), is called 'Interference (through shift)' (Thomason and Kaufman 1988), a term which partially overlaps with the terms 'imposition' (van Coetsem 2000; Winford 2005) and 'substrate influence'.

Contact-induced change can be studied at different time scales, which implies the availability of different data and methods. The following three levels on this temporal continuum may be distinguished:

1. Prehistoric: Some similarities among large numbers of languages in big, often continent-wide, areas can be shown to be the long-term result of multiple, individual contact-induced changes. Details of the historical processes involved must usually remain unclear, including e.g. the sociolinguistic situation and the directionality.

https://doi.org/10.1515/9783110435351-002

2. Reconstructible or attested: Here, individual historical processes can be described for individual pairs of languages, also in terms of directionality. For such cases, proof of contact-induced change requires four components (Thomason 2001: 91–95): (i) the recipient and donor language are or have been in contact; (ii) the recipient and donor language share a feature A; (iii) a pre-contact direct ancestor of the recipient language did not have feature A; and (iv) a pre-contact direct ancestor of the donor language had feature A. Steps (iii) and (iv) can be approximated by reconstruction, if pre-contact ancestor languages are not attested.

3. Ongoing: Here, the creation of new varieties within as little as a few years can be studied, focusing on issues such as matrix language shifts during code-switching, frequency of use of borrowed vs. native material, and the linguistic and sociolinguistic background of speakers involved in producing contact-induced changes.

After a brief discussion of prehistoric contact-induced changes in the following section, the focus in this chapter will be on reconstructible or attested changes, since these kinds of data provide most of the evidence on which inferences regarding asymmetries of contact-induced change can be made.

2. Prehistoric contact-induced change and linguistic universals

At least since Nichols (1992), there is an interest in capturing the interaction between contact-induced change, inheritance, and universality in the world-wide distribution of typological features, such as word order and morphological complexity. Work in this tradition usually starts out by defining macro-areas based on geographical boundaries (Hammarström and Donohue 2014) and historical (including genetic) data, and then compares distributions of features within and across these areas. One method is to study the proportion of families within areas that have biases towards a particular feature, probing the universality of this feature against the probability of areal spread. For example, Bickel (2015) finds strong evidence that language contact favored the occurrence of case categories, such as accusative, in Eurasian languages by showing that their areal spread in Eurasia is independent of the typological correlation between such case marking and verb-final word order. A second example concerns left-branching NPs in Eurasia, which by this method can be shown to be "something that happened to be locally attractive for copying, a mere fashion" (Bickel 2017: 11). Other work in this domain has also revealed asymmetries in the relative borrowability of grammatical subsystems. For instance, classifiers, as one type of nominal classification system, can be shown to be more prone to areal diffusion than genders and noun classes, as another such type (Nichols 1992: 132), and Wichmann and Holman (2009: 54–55) find them to be less stable in terms of inheritance, reflecting the tighter grammatical integration and more abstract meaning of the latter (see Section 3.1.). Borrowed morphological material is usually not identifiable in prehistoric contact-induced change, *Wanderwörter* being interesting exceptions (Haynie et al. 2014).

3. Asymmetries in contact-induced change

3.1. Types of asymmetries

Constraints on contact-induced change are often formulated in terms of asymmetries of borrowability, or 'borrowing hierarchies', such as DERIVATION > INFLECTION, or NOUNS > VERBS. Sometimes authors construe such asymmetries as implicational universals predicting possible vs. impossible constellations in individual languages, e.g. Moravcsik's (1978) claim that no language would borrow bound morphology it if did not also borrow at least one free morpheme. Other authors use such asymmetries to also explain tendencies in a diachronic order, in which different kinds of elements, such as free vs. bound forms, are borrowed (Thomason and Kaufman 1988). Another application of these asymmetries is to describe the relative frequencies of attested cases of borrowing, e.g. free vs. bound forms, in typological samples (Sections 3.2–3.4). Truly universal asymmetries in contact-induced change should surface at all three levels: language-internally, diachronically, and typologically. The theoretical significance of such asymmetries is that they may inform us about properties of human language in terms of the relative stability of linguistic subsystems.

Figure 2.1 summarizes borrowability asymmetries from the literature, focusing on borrowing that includes morphological material (MAT), rather than purely structural

I. FORMAL PROPERTIES OF INDIVIDUAL MORPHEMES
FREE > BOUND
SHARP > FUZZY BOUNDARIES BETWEEN CONSECUTIVE MORPHEMES
TIGHTLY > NOT TIGHTLY INTEGRATED IN MORPHOSYNTACTIC SYSTEM
NO ALLOMORPHS > ALLOMORPHS
OPTIONAL > OBLIGATORY
DERIVATION > INFLECTION
INHERENT INFLECTION > CONTEXTUAL INFLECTION
(Gardani, Arkadiev, and Amiridze 2015)

II. SEMANTIC/PRAGMATIC PROPERTIES OF INDIVIDUAL MORPHEMES
CONCRETE, LEXICAL > ABSTRACT, GRAMMATICAL MEANING
AFFECTIVE > NON-AFFECTIVE MEANING
UNCERTAINTY > CERTAINTY ASSOCIATED WITH MEANING, e.g. FUTURE > PAST
(Matras 2007)

III. PROPERTIES OF SETS OF MORPHEMES
INTERRELATED SETS > NON-INTERRELATED SETS (SEE SECTION 3.4) (Seifart 2012)

IV. INTERSYSTEMIC CONSTRAINTS
GAP-FILLING > NOT GAP-FILLING
CONGRUENT > NON-CONGRUENT STRUCTURES (Field 2002)
PHONOLOGICAL SIMILARITY > DISSIMILARITY

Fig. 2.1: Proposed asymmetries in borrowability
Where not otherwise noted, these are taken from Wilkins (1996), in turn based on earlier work, mainly Weinreich (1953). The symbol '>' means 'more easily borrowable than'

(PAT) borrowing. Properties of individual morphemes (I and II in Figure 1) have been the main focus of research on borrowability, to which most proposed 'borrowing hierarchies' apply. Also included is an asymmetry that applies in situations where sets of morphemes are borrowed (III in Figure 2.1) and 'intersystemic' asymmetries regarding relations between donor and recipient languages (IV in Figure 2.1).

Underlying the compilation of asymmetries in Figure 2.1 is the assumption that each of them has an independent effect on relative borrowability (Wilkins 1996: 110; following Weinreich 1953). However, many of them are not independent, but more or less strongly correlated. For instance, forms with abstract, grammatical meanings are more likely to be bound (affixes) than forms with concrete, lexical meanings. To demonstrate the independence of an individual asymmetry, one would need to keep the values for the remaining asymmetries constant. Attempts at keeping at least some factors constant in typological samples include keeping constant the recipient language (e.g. Romani languages: Elšík and Matras 2006), the donor language (e.g. Spanish: Stolz and Stolz 1996), or bound vs. free morphemes (e.g. affixes only: Seifart 2013, 2017). Another approach is to focus on the relative borrowability of categories that are in quasi-paradigmatic relation, e.g. future vs. past or augmentative vs. diminutive markers (Matras 2007).

The following sections discuss examples of asymmetric borrowing patterns attested in typological samples, focusing on formal and semantic/pragmatic properties of free and bound morphemes (Sections 3.2–3.3), sets of morphemes (Section 3.4), and intersystemic constraints (Section 3.5).

3.2. Asymmetries in borrowing vocabulary

Lexical items are relatively often borrowed. There appear to be few – if any – languages in which no loanwords are in use. A number of asymmetries contribute to this fact: lexical items have concrete, lexical meaning, and are often free. Robust results of relative borrowability within vocabulary are provided by a study that identified loanwords in wordlists of about 1000–2000 entries from 41 languages (Tadmor 2009; Haspelmath and Tadmor 2009). This study establishes a ranking of lexical meanings according to relative borrowability, including a set of 'basic vocabulary' that is relatively resistant to borrowing and that includes as prominent semantic domains body parts, natural phenomena, and some items that relate to human culture, such as HOUSE, NAME, ROPE, and TO TIE. This study also substantiates the NOUN > VERB asymmetry. This asymmetry holds not only in morphologically complex, but also in highly isolating languages, suggesting that it is not reducible to the higher morphological complexity and syntactic integration of verbs, as establishing links with other elements in the clause. In morphologically complex languages, different strategies are employed to incorporate borrowed verb stems, including, but not limited to, inserting bare verb stems and leaving inflection on native auxiliary verbs (Wohlgemuth 2009).

Overall, open-class words are more often borrowed than function words (Tadmor 2009). Within function words, one of the best attested asymmetry concerns conjunctions where the hierarchy BUT > OR > AND has been shown to hold in two independent studies (Stolz and Stolz 1996; Matras 1998). Since most or all other factors are kept constant in comparing the relative borrowability of conjunctions, this empirical finding is particularly informative for modeling contact-induced change (see Section 4).

3.3. Asymmetries in borrowing affixes

Affixes are predicted to be among the least borrowable items by a number of asymmetries, including their bound status, their abstract meaning, and more or less high degree of grammatical integration (see also Chapter 6). Affixes, and in particular derivational affixes, may initially enter a language as part of complex loanwords before they spread to native stems, as was the case with Norman French *-able* in Middle English, for which loanwords from Norman French such as *honourable* served as a 'Trojan horse'. But borrowing inflectional affixes in conjunction with stems is also well attested, resulting in 'parallel systems' (Kossmann 2010), in which entire sets of loanwords are productively inflected with borrowed affixes well beyond isolated instances in learned registers, as with Latin number marking on English *alumnus–alumni*. This raises the question to what extent borrowed affixes are direct effects of language contact, rather than being the result of language-internal spread from borrowed stems to native stems, possibly much later than, and to some extent independent of, the borrowing of the loanwords (Seifart 2015).

With the caveat that the directness of affix borrowing must remain uncertain in many attested cases of affix borrowing, cross-linguistic surveys confirm the asymmetry DERIVATION > INFLECTION (Matras 2007; Seifart 2017). Among the commonly borrowed derivational affixes are nominalizers (especially agent nominalizers), adjectivizers, and diminutives. Among inflectional affixes, those with the most abstract meanings and tight syntactic integration are hardest to borrow, e.g. structural case or subject agreement markers, which are assigned to a word because of the syntactic context in which it appears (Matras 2007; Seifart 2017). Such morphology is called contextual inflection, in contrast with inherent inflection (Booij 1996), which is more easily borrowed. Inherent inflection carries more concrete meaning, and modifies words independently of syntactic context, as in plural markers (Gardani 2012), and is therefore intermediate between derivation and contextual inflection. On the other hand, increasing evidence shows that inflectional affixes, including contextual inflection, are more commonly borrowed than previously assumed (Gardani 2008; Seifart 2017), suggesting that they should be accounted for within models of contact-induced change.

3.4. Asymmetries in sets of borrowed forms

Comparison of languages that borrow entire sets of affixes (or other grammatical morphemes) reveals a tendency towards borrowing interrelated sets of form, e.g. sets consisting of various members of a paradigm of case markers. This results in 'chunks' of borrowed morphology, rather than sets consisting of independent forms, e.g. one case marker, one tense marker, and one number marker (Seifart 2012, 2017; Evans 2016). In such cases of heavy borrowing, the tight, paradigmatic integration of forms, such as case markers or argument indexes, thus favors particular outcomes, even though such forms are correctly predicted to be harder to borrow than other, less tightly integrated forms in the first place.

3.5. Intersystemic asymmetries

Although structural and phonological similarity between donor and recipient language plays a major role in some models of contact-induced change, this has hardly been investigated systematically across large samples of languages. Similarity may facilitate unusual borrowing events, such as borrowing bound affixes, as sometimes observed when closely related languages borrow from each other (Mithun 2013). Structural and phonological similarity due to chance, rather than genealogical relatedness, may also have this effect. For instance, native inflectional suffixes are replaced by phonologically similar ones from Turkish in both Cappadocian Greek and Balkan Romani (Matras 2009: 213–216; Janse 2009), and from Bulgarian in Megleno Romanian (Gardani 2008). Regarding adpositions, according to an early claim (Moravcsik 1978), these can only be borrowed if their position within the adpositional phrase is maintained, i.e. prepositions as prepositions and postpositions as postpositions, and if structural similarity is an absolute constraint on borrowability, e.g. postpositions can only be borrowed by a postpositional language from another postpositional language. Grossman (2014) showed this to be a statistical tendency in a sample of about 100 languages, but also revealed numerous counterexamples of various types.

4. Modeling contact-induced change

Models of contact-induced change aim to capture the interaction between linguistic factors and extralinguistic factors in accounting for attested patterns of contact-induced change. They differ mainly in the relative importance given to sociolinguistic features of 'contact scenarios' vs. features of the grammatical (in particular syntactic) systems of the language involved.

Two basic types of contact scenarios are distinguished, each associated with different patterns of contact-induced change: 'borrowing' and 'interference' scenarios, to use Weinreich's (1953) terms. In the first scenario, 'borrowings' from a donor language are introduced by recipient-language speakers, e.g. English speakers using loanwords from French. In the second one, 'interferences' are introduced by speakers of the donor language into their variety of the recipient language, e.g. French speakers speaking English with a French accent. In Thomason and Kaufman's (1988; further developed in Thomason 2001) model, within the first scenario, there is a continuous range from casual to intense contact, defined by degree of bilingualism and social acceptability of borrowing. Along this continuum, borrowing hierarchies summarized in Figure 2.1 above apply, with casual contact leading to lexical borrowing only (beginning with non-basic vocabulary), and structural borrowing of bound forms occurring only under intense contact.

In the second scenario, contact-induced change is crucially shaped by imperfect learning during second language acquisition, especially of adult learners. This predicts that contact-induced changes will proceed in an almost opposite direction, namely: first affecting phonology (as non-native pronunciation) and grammar (especially syntax, which is often simplified), but not vocabulary, which is relatively easy to acquire, even by adult learners. Accordingly, it has been hypothesized that high-contact languages would be less complex (Kusters 2003; Trudgill 2011), which has been partially substantiated

in quantitative studies showing that languages with more second-language speakers tend to e.g. lose nominal case markers (Bentz and Winter 2013). Pidginization and creole formation (see Chapter 9) are linked to very specific types of contact scenarios involving lack of a common language between e.g. slave populations and plantation owners. These processes are also sometimes argued to involve imperfect acquisition and simplification of lexifier languages, but this is highly contested (Blasi, Michaelis, and Haspelmath 2017). Similarly, mixed languages (Matras and Bakker 2003), e.g. French-Cree Michif, are argued to arise in very specific socio-historical situations as social identity markers through "deliberate manipulation of bilingual repertoires" (Thomason 2008: 51).

Based on van Coetsem's (1988, 2000) work, Winford (2005) has emphasized that the same speakers may be involved in both (imperfectly) acquiring the donor language and borrowing into the donor language from their native language. This necessitates introducing the concepts of dominant language and agentivity, in addition to first and second language, in a model of contact-induced change. How unusual contact-induced changes result from shifting dominance and agentivity over a few decades within a multilingual community has been documented in a community speaking Gurindji and English-based Gurindji Kriol (McConvell and Meakins 2005; Meakins 2011). 'Double agentivity' may also be involved in cases of affix borrowing, in which words consisting of recipient language stems and donor language affixes are first attested in the donor language (e.g. Quechua-Spanish *warmin-ero* 'woman-izer' used in Spanish) before they were used by (possibly the same) speakers in the recipient language, resulting in borrowed affixes (Seifart 2015).

Another type of sociolinguistic factor are cultural constraints on lexical borrowing. For instance, the linguistic exogamy (i.e. marrying outside of speech community only) practiced by some societies of the Northwest Amazon requires maintenance of vocabulary as social identity marker, but also implies heavy multilingualism. This explains how remarkably low rates of lexical borrowing can coexist with rampant contact-induced structural changes, and in some cases, borrowed affixes (Epps and Michael 2017).

Other models emphasize that contact-induced changes should be in accordance with recipient language speakers' grammatical, and in particular syntactic, representations of their language (Poplack 2009; Silva-Corvalán 1994; Myers-Scotton 2002; Field 2002). Within these approaches, for instance, Myers-Scotton's (2002) 4-M(orpheme-type) model predicts a decreasing likelihood of code-switching, and subsequent borrowing, of different morpheme types, which is largely consistent with the asymmetry DERIVATION > INHERENT INFLECTION, and an absolute constraint on borrowing contextual inflection. Within this framework, the difficulty or impossibility of borrowing certain morpheme types is related to their increasing integration in syntactic structures, modeled in terms of universal grammar. An important determinant of contact-induced change in these approaches is the distance between the donor and recipient language ('intersystemic asymmetries' in Figure 2.1), where it is often assumed that there are absolute constraints to the effect that code-switching, and consequently borrowing, is only permitted where two languages share a particular structure, e.g. constituent order. Such constraints have been increasingly challenged by empirical data, for instance, showing the possibility of borrowing prepositions as such in postpositional languages (Grossman 2014).

Focusing on the goal-oriented nature of communication, Matras (2007, 2009) provides a model for contact-induced change which, among other things, accounts for a number of asymmetries by the 'selection malfunction hypothesis'. According to this

hypothesis, borrowing is more likely for forms that express uncertainty and related meanings, such as future tense. This uncertainty potentially challenges speakers' authority and compromises their control over the selection of languages in their repertoire, and thus leads to code-switching and, eventually, borrowing. This approach is able to explain the asymmetry BUT > OR > AND (where BUT expresses a potentially challenging adversative relation), and others otherwise unaccounted for.

Finally, Muysken (2013) has shown how many of the considerations discussed so far can be accommodated in a single model as optimality-theoretical constraints related to four 'bilingual optimization strategies': (i) borrow into first language only where necessary; (ii) use as much as possible of the second language; (iii) produce structures and words which share properties of the first and second language; and (iv) use universal combinatory principles to combine fragments from different languages independently of the grammars involved. These four constraints can then be variably ranked according to the concrete contact situation to correctly predict the actual outcome as e.g. favoring or disfavoring borrowing of various kinds.

5. Conclusions and outlook

Increasing empirical evidence of contact-induced change in the world's languages challenges the existence of absolute constraints on contact-induced change, but promises to uncover significant asymmetries in the form of statistical tendencies. One challenge for modeling contact-induced change is to test the separate effects of the individual factors involved, and capture their interactions and dependencies with universal properties of human language and language genealogy. Progress in this respect will require scaling up the descriptive-empirical groundwork and submitting the enlarged data sets to large-scale statistical testing. Four areas appear particularly promising for future work: (i) text frequencies (van Hout and Muysken 1994) and other evidence gleaned from corpora documenting naturally occurring multilingual speech (Adamou 2016); (ii) the shifting agency and dominance of speakers in multilingual communities, as observable in sociolinguistically contextualized longitudinal studies (McConvell and Meakins 2005); (iii) asymmetries in the relative stability of grammatical features, as observable in large typological databases (Wichmann and Holman 2009); and (iv) the use of genetic evidence to infer language shift and contact scenarios to provide insights into prehistoric processes that may have had an impact on contact-induced language change (Pakendorf 2014).

6. References

Adamou, Evangelia
2016 *A corpus-driven approach to language contact: Endangered languages in a comparative perspective*. Berlin & Boston: De Gruyter Mouton.

Bentz, Christian & Bodo Winter
2013 Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change* 3(1). 1–27. doi:10.1163/22105832–13030105.

Bickel, Balthasar
2015 Distributional typology: Statistical inquiries into the dynamics of linguistic diversity. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis, 2nd edn.*, 901–923. Oxford: Oxford University Press.
Bickel, Balthasar
2017 Areas and universals. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 40–54. Cambridge: Cambridge University Press.
Blasi, Damián E., Susanne Maria Michaelis & Martin Haspelmath
2017 Grammars are robustly transmitted even during the emergence of creole languages. *Nature Human Behaviour* s41562–017–0192–4. doi:10.1038/s41562–017–0192–4.
Booij, Geert E.
1996 Inherent versus contextual inflection and the split morphology hypothesis. In Geert Booij & Jaap van Maarle (eds.), *Yearbook of morphology 1995*, 1–15. Dordrecht: Kluwer Academic Publishers.
Elšík, Viktor & Yaron Matras
2006 *Markedness and language change: The Romani sample*. Berlin: Mouton de Gruyter.
Epps, Patience & Lev Michael
2017 The areal linguistics of Amazonia. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 934–963. Cambridge: Cambridge University Press.
Evans, Nicholas
2016 As intimate as it gets? Paradigm borrowing in Marrku and its implications for the emergence of mixed languages. In Felicity Meakins & Carmel O'Shannessy (eds.), *Loss and renewal: Australian languages since colonisation*, 29–56. Boston & Berlin: De Gruyter Mouton.
Field, Frederic W.
2002 *Linguistic borrowing in bilingual contexts*. Amsterdam & Philadelphia: John Benjamins.
Gardani, Francesco
2008 *Borrowing of inflectional morphemes in language contact*. Frankfurt/Main: Peter Lang.
Gardani, Francesco
2012 Plural across inflection and derivation, fusion and agglutination. In Lars Johanson & Martine Robbeets (eds.), *Copies versus cognates in bound morphology*, 71–97. Leiden & Boston: Brill.
Gardani, Francesco, Peter Arkadiev & Nino Amiridze
2015 Borrowed morphology: An overview. In Francesco Gardani, Peter Arkadiev & Nino Amiridze (eds.), *Borrowed morphology*, 1–23. Berlin & Boston: De Gruyter Mouton.
Grossman, Eitan
2014 Evaluating a universal of borrowing: Linear order preservation. Paper given at University of Zurich, Department of Comparative Linguistics Forschungskolloquium, 23 May 2014.
Hammarström, Harald & Mark Donohue
2014 Some principles on the use of macro-areas in typological comparison. *Language Dynamics and Change* 4(1). 167–187. doi:10.1163/22105832–00401001.
Haspelmath, Martin & Uri Tadmor (eds.)
2009 *The World Loanword Database (WOLD)*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wold.clld.org/.
Haynie, Hannah, Claire Bowern, Patience Epps, Jane Hill & Patrick McConvell
2014 Wanderwörter in languages of the Americas and Australia. *Ampersand* 1 (Supplement C). 1–18. doi:10.1016/j.amper.2014.10.001.
Janse, Mark
2009 Greek-Turkish language contact in Asia Minor. *Études Helléniques/Hellenic Studies* 17(1). 37–54.

Johanson, Lars
2008 Remodeling grammar: Copying, conventionalization, grammaticalization. In Peter Siemund & Noemi Kintana (eds.), *Language contact and contact languages*, 61–79. Amsterdam & Philadelphia: John Benjamins.
Kossmann, Maarten
2010 Parallel system borrowing: Parallel morphological systems due to the borrowing of paradigms. *Diachronica* 27(3). 459–487.
Kusters, Wouter
2003 *Linguistic complexity: The influence of social change on verbal inflection*. Utrecht: LOT.
Matras, Yaron
1998 Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36(2). 281–331.
Matras, Yaron
2007 The borrowability of structural categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 31–73. Berlin & New York: Mouton de Gruyter.
Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.
Matras, Yaron & Jeanette Sakel (eds.)
2007 *Grammatical borrowing in cross-linguistic perspective*. Berlin & New York: Mouton de Gruyter.
Matras, Yaron & Peter Bakker (eds.)
2003 *The mixed language debate: Theoretical and empirical advances*. Berlin: Mouton de Gruyter.
McConvell, Patrick & Felicity Meakins
2005 Gurindji Kriol: A mixed language emerges from code-switching. *Australian Journal of Linguistics* 25(1). 9–30.
Meakins, Felicity
2011 Borrowing contextual inflection: Evidence from northern Australia. *Morphology* 21(1). 57–87. doi:10.1007/s11525–010–9163–4.
Mithun, Marianne
2013 Challenges and benefits of contact among relatives: Morphological copying. *Journal of Language Contact* 6(2). 243–270. doi:10.1163/19552629–00602003.
Moravcsik, Edith A.
1978 Language contact. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.), *Universals of human language. Vol. 1: Method & Theory*, 94–122. Stanford: Stanford University Press.
Muysken, Pieter
2013 Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16(04). 709–730. doi:10.1017/S1366728912000727.
Myers-Scotton, Carol
2002 *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
Nichols, Johanna
1992 *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
Pakendorf, Brigitte
2014 Historical linguistics and molecular anthropology. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 627–641. Oxon & New York: Routledge.
Poplack, Shana
2009 Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics* 18(7–8). 581–618. doi:10.1515/ling.1980.18.7–8.581.

Seifart, Frank
2012 The principle of morphosyntactic subsystem integrity in language contact: Evidence from morphological borrowing in Resígaro (Arawakan). *Diachronica* 29(4). 471–504.
Seifart, Frank
2013 *AfBo: A world-wide survey of affix borrowing*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://afbo.info.
Seifart, Frank
2015 Direct and indirect affix borrowing. *Language* 91(3). 511–532. doi:10.1353/lan.2015.0044.
Seifart, Frank
2017 Patterns of affix borrowing in a sample of 100 languages. *Journal of Historical Linguistics* 7(3). 389–431. doi:10.1075/jhl.16002.sei.
Silva-Corvalán, Carmen
1994 *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
Stolz, Christel & Thomas Stolz
1996 Transpazifische Entlehnungsisoglossen: Hispanismen in Funktionswortinventaren beiderseits der Datumsgrenze. In Norbert Boretzky, Werner Enninger & Thomas Stolz (eds.), *Areale, Kontakte, Dialekte. Sprache und ihre Dynamik in mehrsprachigen Situationen. Beiträge zum 10. Bochum-Essener-Symposium vom 30. 06.–01. 07. 1995 an der Universität GH Essen*, 262–291. Bochum: Universitätsverlag Dr. N. Brockmeyer.
Tadmor, Uri
2009 Loanwords in the world's languages: Findings and results. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 55–75. Berlin: Mouton de Gruyter.
Thomason, Sarah G.
2001 *Language contact: An introduction*. Washington, D.C.: Georgetown University Press.
Thomason, Sarah G.
2008 Social and linguistic factors as predictors of contact-induced change. *Journal of Language Contact* 2(1). 42–56. doi:10.1163/000000008792525381.
Thomason, Sarah G. & Terrence Kaufman
1988 *Language contact, creolization, and genetic linguistics*. Berkeley, CA: University of California Press.
Trudgill, Peter
2011 *Sociolinguistic typology: Social determinants of linguistic complexity.* Oxford: Oxford University Press.
van Coetsem, Frans
1988 *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
van Coetsem, Frans
2000 *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
van Hout, Roeland & Pieter Muysken
1994 Modeling lexical borrowability. *Language Variation and Change* 6(1). 39–62. doi:10.1017/S0954394500001575.
Weinreich, Uriel
1953 *Languages in contact*. New York: Linguistic Circle of New York.
Wichmann, Søren & Eric W. Holman
2009 *Temporal stability of linguistic typological features*. Munich: LINCOM EUROPA.
Wilkins, David P.
1996 Morphology. In Hans Goebl, Peter H. Nelde, Zdeněk Starý & Wolfgang Wölck (eds.), *Contact linguistics: An international handbook of contemporary research*. Vol. 1, 109–117. Berlin & New York: Mouton de Gruyter.

Winford, Donald
2005 Contact-induced changes. Classification and processes. *Diachronica* 22(2). 373–427.
Wohlgemuth, Jan
2009 *A typology of verbal borrowings*. Berlin & Boston: De Gruyter Mouton.

Frank Seifart, Amsterdam (The Netherlands) and Cologne (Germany)

3. Universal processes in contact-induced syntactic change

1. Introduction
2. Research context
3. Data and framework
4. Contact-driven syntactic changes
5. References

1. Introduction

Contact-induced syntactic change represents a key topic within contact linguistics. However, due to the lack of systematicity in terminology, quite dissimilar contact-driven processes leading to syntactic change have often been classified under the same label. Terms such as transfer, grammatical borrowing, substrate effect, etc. have frequently been used interchangeably in the literature. The lack of consensus on the origin of observable contact-induced syntactic changes is in part due to this terminological confusion.

Section 2 of this chapter first offers an overview of some of the main claims on the nature of contact-induced syntactic change by relying on a number of examples drawn from several contact varieties. Subsequently, in Section 3, I focus on a narrower set of languages, the Afro-Hispanic languages of the Americas (AHLAs), for which certain universal processes are described from a generative standpoint. The phenomena analyzed, I argue, are not necessarily ascribable to the specific superstrate and substrate languages (i.e. Spanish, Bantu, etc.), rather, they are largely due to the modular architecture of the language faculty, and thus universal. Section 4 offers an account of some universal contact-driven processes leading to syntactic change.

2. Research context

Different authors hold divergent perspectives on the nature of contact-induced syntactic change. Thomason and Kaufman (1988: 74–76), for example, claim that given the right sociolinguistic context almost anything – including syntax – can be 'borrowed'. They provide a borrowing scale, which ranges from basic lexical borrowings – occurring in

https://doi.org/10.1515/9783110435351-003

cases of casual contact − to heavy structural borrowings, which take place when different languages intensively coexist for prolonged periods of time (Haugen 1950; Muysken 1981; Whitney 1881 on 'hierarchies of borrowability'). In this respect, Harris and Campbell (1995) appear to be of the same opinion. Indeed, in their well-known book on historical syntax, they echo Thomason and Kaufman (1988) by stating that, "given enough time and intensity of contact, virtually anything can (ultimately) be borrowed" (Harris and Campbell 1995: 149).

In their extensive review of the literature on contact-induced syntactic change, Harris and Campbell (1995: 120−150) point out a series of counter-examples to all the main theories of structural transfer. For example, to show how syntactic borrowing would not be limited to cases where "the two languages had a good deal of syntactic similarity to begin with", as hypothesized by Allen (1980: 380; see also Meillet 1914: 86, 87; Weinreich 1953: 25), they present examples from American Finnish, which has borrowed syntactic patterns from English, even though they are not at all compatible with Finnish syntax. Indeed, they show how Finnish does not allow infinitival complements of nouns or adjectives with subject rising, a pattern found in English and, consequently, also in American Finnish (Harris and Campbell 1995: 125).

(1) American Finnish
Tämä oli ensimmäinen kerta mei-lle mennä tä-llä laiva-lla.
this was first time us-for to go this-on ship-on
'This was the first time we went on this ship.'

(2) Standard Finnish
Tämä oli ensimmäinen kerta, kun oli-mme matkusta-neet tä-llä
this was first time when had-we travel-PAST.PTCPL this-on
laiva-lla.
ship-on
'This was the first time we went on this ship.'

Others, such as Winford (2003), have been more cautious with examples of this type. In Winford's view, it is not so easy to accept that everything can be *directly* borrowed. Indeed, he believes that the presence of a grammatical structure in one linguistic system proceeding from another language is almost always *mediated* by other processes, such as lexical borrowing. He states that "despite the many claims concerning the 'borrowing' of phonology, morphology, and *even syntax* [emphasis mine], there has been no convincing demonstration that such structural change occurs without mediation by some other medium or process" (Winford 2003: 61).

The reason behind such a different perspective on the issue is, in part, a matter of terminology. In Winford's view, the word 'borrowing', as used by authors such as Thomason and Kaufman (1988), would be an overly general umbrella term, which refers to at least two very different cognitive processes. Winford (2003) adopts van Coetsem's (1988) notion of language agentivity, which assumes the existence of two different types of transfer: a) borrowing, which primarily affects lexical items and occurs under Recipient Language (RL) agentivity; and b) imposition, which concerns structure (syntax and phonology) and is driven by Source Language (SL) agentivity. Thus, according to this view, it seems more plausible that a given speaker, dominant in a given language (lan-

guage X), imposes X structures on the language he is trying to speak (language Y), in which he is not necessarily proficient, rather than borrowing Y syntax and/or phonology into his own language.

Winford (2003: 69) exemplifies the concepts of 'structural transfer through lexical mediation' (RL agentivity) and 'imposition' (SL agentivity) by providing King's (2000: 139–149) Prince Edward Island French data. In a situation of unstable bilingualism, English preposition stranding would have entered Prince Edward Island French, thus, producing the structural changes observable in (3), which significantly contrast with the standard French patterns in (4).

(3) Prince Edward Island French
Quoi ce-qu'ils parlent?
what that-they are-talking about
'What are they talking about?'

(4) Standard French
De quoi parlent-ils?
about what are-talking-they
'What are they talking about?'

Winford (2003) builds on King's (2000: 136) claim that "there is a close relationship between borrowing prepositions and having preposition stranding," and highlights how the intense borrowing of English prepositions would have caused a reconfiguration of the Prince Edward Island French prepositional system, thus showing that what may appear to be a case of direct structural borrowing "is actually an instance of mediated lexical borrowing" (Winford 2003: 69), under RL agentivity. Moreover, Winford (2003) highlights how Prince Edward Island French would be particularly at risk due to intense pressure from English, which is becoming the dominant language in certain local communities. Such a process, therefore, would also imply language attrition and shift, which promotes structural transfer via SL agentivity (i.e. imposition) from the dominant language on the minority one.

While the aforementioned models are primarily related to cases of language maintenance and shift, another facet of contact-induced syntactic change is concerned with the creation of new languages, such as pidgins, creoles and mixed languages (Velupillai 2015). In this respect, it is worth highlighting the so-called 'Creole Debate' (McWhorter 2018), which focuses on the structural and typological status of creole languages. In fact, among linguists there is no common consensus on how creole languages are to be defined, nor on how such contact varieties should be classified. Some authors have tried to classify creoles according to their structural properties (Bickerton 1981) or as a typological class (Bakker et al. 2011; McWhorter 1998; Seuren and Wekker 1986). Others have adopted a comparative structural approach, which categorizes these vernaculars according to the relative distance from their lexifiers, thus leading to more terms, such as 'semi-creoles' (Holm 1992; Schneider 1990) or 'intermediate creoles' (Winford 2000). Other linguists have rejected these analyses and prefer to depict creoles as byproducts of their shared sociocultural history, often related to slavery and plantation societies (DeGraff 2003; Mufwene 1997), thus claiming that they show *nothing exceptional* from a strictly linguistic point of view. A recent contribution to this debate is the one offered by Aboh (2015), who conceives of these languages as mixed grammars. In his view, a

creole, like any other contact variety, would be the result of a mix of features proceeding from the pool of languages that were in contact, according to a competitive mechanism driven by environmental and/or ecological factors (Mufwene 2001).

Aboh (2015: 10) illustrates how this mixing mechanism would work by presenting examples of noun phrases in Sranan, an English-based creole with Gbe as the main substrate. The author claims that in this creole, like in Gungbe, which belongs to the Gbe languages, nouns appear bare (5) and may co-occur with discourse-specific markers taking on either definite, indefinite, or generic readings, depending on context (6).

(5) Sranan

a. *Kofi, go na wowoyo go bai bana tya kon gi mi.*
Kofi go LOC market go buy banana carry come give me
'Kofi, go to the market to buy me (plantain) banana(s).'

b. *Kofi, teki a bana tya kon gi mi.*
Kofi take DET banana carry come give me
'Kofi, give me the (plantain) banana (in question).'

(6) Gungbe

a. *Àxwlé nɔ dù gbàdó.*
turtle-dove HAB eat corn
'A turtle-dove habitually eats corn / Turtle-doves habitually eat corn.'

b. *Àxwlé dù gbàdó lɔ.*
turtle-dove eat corn DET
'A turtle-dove ate the corn (in question) / Turtle-doves ate the corn (in question).'

Aboh (2015: 8) argues that creoles are not at all "exceptional", rather, they "represent a normal instance of language change resulting from the contact between typologically different and genetically unrelated languages (e.g. Romance/German vs. Kwa/Bantu [Niger-Congo])". This would be the only reason why the structural changes observed in creoles tend to be more contrastive than in other contact varieties. This idea is compatible with Blasi, Michaelis and Haspelmath's (2017: 723) proposal that "grammars are robustly transmitted even during the emergence of creole languages" (see McWhorter [2018] for a drastically divergent opinion).

Turning now from this overview of some of the main models and debates characterizing the literature on contact-induced syntactic change (see Harris and Campbell 1995: 120–150; Muysken 1996: 119–122; and Roberts 2007: 383–442 for additional data), I focus on another set of contact-driven syntactic phenomena, which tend to occur when two or more linguistic systems interact, independent of the languages involved. The mechanisms behind such phenomena, I argue, are universal. They are a byproduct of the same second-language acquisition strategies, which becomes particularly evident when it comes to analyzing constructions implying high processing demands on modular interfaces, i.e. the syntax/discourse interface, the morphology/semantics interface, etc. (White 2011). To exemplify such phenomena, I use data from my own fieldwork (Sessarego 2013b, 2014, 2015) on three AHLAs: Yungueño Spanish (Bolivia), Chinchano Spanish (Peru) and Chota Valley Spanish (Ecuador).

3. Data and framework

The AHLAs, the languages that developed in Latin America from the contact of African languages and Spanish in colonial times, show grammatical peculiarities that deviate from standard Spanish and which have repeatedly been reported in the literature as indicators of a previous creole stage. Some of these recurring properties are: (7) presence of bare nouns in argument position; (8) variable number and gender agreement across the Determiner Phrase (DP); (9) variable subject-verb agreement; (10) use of non-emphatic, non-contrastive overt subjects; (11) lack of subject-verb inversion in questions (a feature that is also common in Caribbean Spanish, for which a creole ancestor has been proposed in several cases; see Otheguy 1973).

(7) Yungueño Spanish
Chancho come papa.
pig-SG eat potato-SG
'Pig(s) eat potato(es).'
(Sessarego 2014: 70)

(8) Yungueño Spanish
Todo las casa bonito.
all-M-SG the-F-PL house-F-SG nice-M-SG
'All the nice houses.'
(Sessarego 2014: 131)

(9) Chota Valley Spanish
Ellos dijo, ellos dijeron eso.
they said-3-SG they said-3-PL it
'They said, they said so.'
(Sessarego 2013b: 76)

(10) Chinchano Spanish
Yo digo que yo lo merezco porque yo eso trabajé.
I say that I CL deserve because I this worked
'I say that I deserve it because I worked for this.'
(Sessarego and Gutiérrez-Rexach 2017: 58–59)

(11) Chinchano Spanish
¿Qué ella dijo?
what she said
'What did she say?'
(Sessarego 2015: 58)

Traditionally, the presence of the aforementioned properties in a number of AHLAs has been analyzed as evidence suggesting that all these contact varieties would have derived from a single creole (Monogenesis Hypothesis) and that, in a subsequent phase, would have gradually converged toward more standard varieties of Spanish (Decreolization Hypothesis) (de Granda 1968). My version of the story is quite different. I maintain that these properties should be analyzed as result of advanced universal second-language acquisition strategies, which does not imply any previous (de)creolization phase. In par-

ticular, these grammatical phenomena are byproducts of processability constraints affecting language interfaces. They are not only peculiar to the AHLAs, but rather appear to be quite common to all cases of language contact (Sessarego 2013a).

The model proposed here to analyze these grammatical elements is built on recent formal hypotheses on the architecture of the language faculty. Jackendoff's (1997) framework on linguistic interfaces appears to be well-suited for this purpose, since it embraces a modular conception of the language faculty and envisions a parallel interface-based interaction among the different modules. In recent years, research on language development has focused on the so-called 'Interface Hypothesis' (e.g. Sorace 2004; Tsimpli et al. 2004). All of these studies essentially maintain the core idea that certain constructions involving high processing demands on the interface between different linguistic modules may be more difficult to master in second-language acquisition and are the first to be eroded in first-language attrition. Along these lines of reasoning, I analyze the morpho-syntactic properties commonly ascribed to a previous creole phase for the AHLAs as result of *advanced conventionalized second-language acquisition strategies*.

I posit that the AHLAs may be analyzed as byproducts of first-language acquisition (nativization) of advanced second-language (L2) grammars. My thoughts regarding Yungueño Spanish, Chota Valley Spanish, and Chinchano Spanish (and most of the other Afro-Hispanic varieties) are that African slaves in colonial times managed to acquire relatively close approximations to Spanish (the target language, TL). This means that each individual internalized an L2 grammar out of a set of possible grammars. Their linguistic output (x, y, z) represented the Primary Language Data (PLD) for the subsequent generations of speakers, who acquired that vernacular as their L1 and conventionalized it at the community level. The aforementioned cross-generational language transmission is represented in (12), in which G1 and G2 are two possible grammars with different structural configurations (see Roberts [2007: 384–391] for a similar yet different model of contact-induced cross-generational syntactic change).

(12) a. Individual from Generation 1:
TLy → UG driving L2 acquisition → G1 → set of outputs

b. Individual from Generation 2:
PLDx → UG driving L1 acquisition → G2 → set of outputs Z

The schema in (12b) depicts a case of nativization, consisting of Generation 2's L1 acquisition. This translates to an L1 grammar (G2) built on L2 inputs. Consequently, G2 shows crystallized aspects of an L2, which are acquired as an L1.

4. Contract-driven syntactic changes

4.1. Bare nouns

In recent studies, I have offered an account of bare nouns in the AHLAs (Sessarego 2013b, 2014, 2015). There, I tested Chierchia's (1998) Nominal Mapping Parameter (NMP) and speculated on the nature of the nominal domain of these vernaculars in light

of Longobardi's (1994) generalization on the structure of DPs. The findings suggest that the NMP does not adequately explain the data. Indeed, these dialects do not fit into Chierchia's (1998) typological classification because: they present bare nouns in argument position, see (7); they display plural morphology on D categories, see (8); and the count/mass distinction is instantiated at the lexical level, see (13).

(13) Yungueño Spanish

a. *Mucho vaso tomó Pablo.*
much glass drank Pablo
'Pablo drank many glasses.'
(Sessarego 2014: 70)

b. *Mucho vino tomó Pablo.*
much wine drank Pablo
'Pablo drank much wine.'
(Sessarego 2014: 70)

In these varieties, bare nouns lack number specification. Number and quantificational force are encoded at the D level. This particular configuration allows for a wide array of interpretive possibilities. In general, these dialects instantiate a very flexible system, where the absence of overt morphological marking on nouns allows for contextual factors to determine the interpretation of bare DPs.

These characteristics make the AHLA nominal domain quite similar to the one described by Aboh (2015) for Sranan, even though the Gbe family does not appear to have had any significant effect on the formation of these languages. In fact, bare nouns of this type can commonly be detected in a variety of interlanguage grammars, even for very advanced levels of proficiency (García Mayo and Hawkins 2009). Thus, common SLA processes, rather than a specific substrate effect, may be able to account for the phenomena detected in all of these contact varieties. As far as Spanish is concerned, these phenomena have been well-documented in a number of cases (Montrul and Ionin 2010; Sánchez and Giménez 1998). In recent years, the presence of bare nouns in interlanguage grammars has been explained as a result of a syntax/semantics interface mismatch, which tends to persist at advanced stages of L2 acquisition (Ionin and Montrul 2010), and therefore should not necessarily be taken as a sign of AHLA (de)creolization.

4.2. Nominal and verbal morphology

As for nominal and verbal morphology (Examples 8–9), generative linguistics has often considered the morphological richness of Spanish as a potential explanation for N-to-Num movement and V-to-Infl movement, among other phenomena (Picallo 1991; Pollock 1989). Nevertheless, the AHLAs, deprived of such morphological characteristics, show exactly the same noun+adjective and verb+adverb order combinations of standard Spanish. This indicates that agreement, at least in these clear cases, cannot be the trigger of movement, which may be driven by other mechanisms, such as the Extended Projection Principle (EPP) (Chomsky 1982) or categorical features.

As for these variable instances of phi-features (i.e. gender, number, person), cross-dialectal differences between the AHLAs and standard Spanish can be accounted for in a systematic fashion, as computationally determined by differences in the specification of lexical and functional items and by restrictions on syntactic operations (Sessarego 2013c). Agreement variability can be formalized in line with the model proposed by Adger and Smith (2005), who argue for characterizing syntactic variation in terms of (un)interpretable features. According to this proposal, certain uninterpretable features may be present or active in one category or element, but absent or inactive in another. Since they are uninterpretable, they would have no semantic repercussions at Logical Form (LF). As a way of exemplifying one of these processes found across the AHLA nominal and verbal domains, (14) depicts variable number agreement across the Chota Valley Spanish DP. As can be seen, in contrast with the agreeing items, the elements showing single default forms do not carry any Num feature specification.

(14) Chota Valley Spanish

a. [DP Mucho [NumP [nP [NP hermano]]]]
*i*Num[PL]

b. [DP Muchos [NumP [nP [NP hermano]]]]
*u*Num[PL] *i*Num[PL]

c. [DP Muchos [NumP [nP [NP hermanos]]]]
*u*Num[PL] *i*Num[PL] *u*Num[PL]

much-SG./PL. brother-SG./PL.
'Many brothers'

(Sessarego 2013b: 70).

Variable phi-agreement is a well-known characteristic of L2 Spanish varieties (Montrul 2004), which − as far as the nominal domain is concerned − has been systematically documented for both gender and number features, even in advanced interlanguages (Bruhn de Garavito and White 2000; Franceschina 2005; White et al. 2004). L2 impoverished agreement has often been linked to processing constraints applied at the morphology-semantics interface (White 2011: 585−586). This could be due to the difficulty of mapping semantic concepts onto overt grammatical morphemes in SLA. For this reason, the development of morphological agreement in the L2 is seen by some researchers as the 'bottleneck' of acquisition (Slabakova 2008). This is true not only developmentally, but also at the end-state of L2 acquisition, thus leading to fossilization phenomena consisting of significant morphological reductions (Lardiere 2007).

Subject-verb agreement mismatches, often resulting in default 3rd person singular verb forms, are well-attested in L2 Spanish speech (Montrul 2004: 146−152). Default forms of this type also tend to persist in the speech of students learning Spanish in formal settings, such as in foreign language classrooms (Slabakova 2009). As for the formalization of these phenomena in the AHLAs (Example 11), it may be postulated that among the available lexical items of these Afro-Hispanic vernaculars there are two different and competing tense heads (T1 and T2). T1 would carry a full set of features (number, person, tense, case), like in standard Spanish, while T2 would not be endowed with the number and person specifications. The application of the operation *Agree* between T1 and the subject pronoun would lead to a standard-like agreement configuration,

as illustrated in (15–16). Conversely, a default 3rd person singular verb form would result from the lexical selection of T2, as formalized in (17–18).

(15) T1 [tense:past, *u*case:nom, *u*num:, *u*pers:] pronoun [num:pl, pers:3, case:nom]
→ T1 [tense:past, num:pl, pers:3] pronoun [num:pl, pers:3, case:nom]

(16) Spell-Out: *Ellos* *dijeron*
they-NOM say-PAST.3.PL

(17) T2 [tense:past, *u*case:nom] pronoun [num:pl, pers:3, *u*case:]
→ T2 [tense:past] pronoun [num:pl, pers:3, case:nom]

(18) Spell-Out: *Ellos* *dijo*
they-NOM say-PAST.3.PL

Thus, if we analyze the cases of phi-agreement mismatches found in the AHLAs in light of recent proposals on the nature of the *morphology-semantics interface* in second-language acquisition (Slabakova 2008; White 2011), it is not difficult to picture how certain advanced L2 patterns may have been conventionalized at the community level in a social context in which acquisition proceeded in a naturalistic and/or untutored way. The properties exemplified in (8–9) are, therefore, common phenomena related to the difficulty of mapping semantic concepts onto L2 morphemes (i.e. the 'bottleneck' of second-language acquisition). Given the universal mechanisms underlying these processes and their consequent generalized presence in contact grammars, morphological reductions of this type should not be seen as just 'creole-like' phenomena.

4.3. Overt subject pronouns and subject-verb inversion in questions

The impoverished verbal agreement system characterizing the AHLAs has indirectly contributed to the development and stabilization in these vernaculars of non-emphatic, non-contrastive overt subject pronouns, which, by encapsulating person and number features, compensate for the morphological reduction of Infl. In addition, the presence of these grammatical elements in those dialects is also explainable as result of second-language acquisition processing constraints applying at the syntax/pragmatics interface (Sorace and Serratrice 2009), since both structural and discourse-related features are involved. Indeed, null subjects (*pro*) are generally associated with old information and are used in topic-continuity and non-contrastive constructions; thus, their proficient use requires considerable cognitive workload on both the syntactic and the pragmatic modules of the language faculty.

A formalization of the difference between an overt subject pronoun and *pro* in a Null Subject Parameter (NSP) language (Rizzi 1982) has been provided by Grimshaw and Samek-Lodovici (1998), who suggest that the overt category would be endowed with a [+topic shift] feature, which would not be present in null subjects. The use of overt pronouns in the AHLAs in contexts in which *pro* would sound more pragmatically felicitous for other native varieties of Spanish (Example 10) appears to suggest that in these

dialects such categories are not always systematically associated with [+topic shift] features. For this reason, it may be claimed that two different types of overt pronouns are available in the lexicon of these Afro-Hispanic vernaculars: type 1 would be endowed with [+topic shift] features (like in standard Spanish); type 2 would not carry such a specification, as in many advanced L2 and heritage varieties of Spanish (19).

(19) ***Yo*** *digo que* ***yo*** *lo merezco porque* ***yo*** *eso trabajé.*
[+topic shift] [−topic shift] [−topic shift]
'I say that I deserve it because I worked for it.'

A number of SLA studies over the past two decades have explored the vulnerability of the syntax-pragmatics interface; in particular, when it comes to the use of subject pronouns (Sorace 2004). All report that native-like proficiency in this aspect of the grammar is particularly hard to obtain, even at very advanced stages of acquisition. The use of non-emphatic, non-contrastive subject pronouns has also been linked to the presence of non-inverted questions (Example 11) in certain Romance varieties (Camacho 2013).

The formal mechanisms underlying the L2 acquisition of *wh*-fronting and subject-verb inversion in questions have long-been a topic of research in generative syntax (Birdsong 1992; White and Juffs 1998). Results indicate that obtaining native-like mastery of these constructions may be challenging, especially if the L2 speakers proceed from an L1 background lacking these operations (e.g. in Chinese, Korean, and Japanese). While certain formal analyses propose that the presence/absence of subject-verb inversion could be related to the existence of a parametrized interrogative feature strength ([+wh/Q]) in the Complementizer head, which would be hard to master even at advanced stages of L2 acquisition (Cuza 2013; Guerra Rivera, Coopmans, and Baauw 2015), linguistic models that do not necessarily embrace the parametric/minimalist framework do acknowledge that subject-verb inversion is only mastered at advanced L2 levels of proficiency (Pienemann 1998). This is another piece of evidence indirectly indicating that this property − along with all the phenomena exemplified in (7−11) − is the result of universal second-language acquisition strategies, which should not necessarily be ascribed to either (de)creolization or any specific substrate effect.

These considerations, however, are not meant to indicate that the L1 does not affect the acquisition of an L2; interference effects certainly play a key role in creole formation and in language contact in general (Lefevbre, White, and Jourdan 2006; Kouwenberg and Singler 2008). That being said, the phenomena we just analyzed do not appear to be primarily driven by any specific L1 effect, but rather by universal processing constraints affecting linguistic interfaces.

During the past few decades, the literature on SLA and interfaces has seen a proliferation of studies (see Rothman and Slabakova 2011 for a review). A number of authors have proposed models to explain the nature of mapping between language modules and the inherent challenges that the L2 learner faces when dealing with different interfaces (independently of the L1s and the target languages involved). While some researchers maintain that the most difficult phenomena to master are those that imply high processing demands on external interfaces (i.e. pronominal use, related to the syntax-pragmatics interface [Sorace and Serratrice 2009]), Slabakova's (2008) Bottleneck Hypothesis ascribes most of the problematic issues to the internal interface between morphology and semantics. Independent of whichever hypothesis is the most suitable to account for lan-

guage development, it is clear that the phenomena investigated in this chapter represent challenging structures for all L2 speakers, thus they are likely to be detected in any contact variety.

Finally, it must be pointed out that, in all cases of language contact, structural and processing factors are not the only relevant ones for a complete analysis (see Myers-Scotton's 2002 code-switching model); rather, social and cultural practices are also key to understanding contact-induced language change (see Chapter 2), such as the major role played by exogamic marriages in explaining the high degree of structural diffusion across Amazonian languages (Epps and Michael 2017), or the significance of hybrid social identity in explaining the emergence of mixed languages like Michif and Media Lengua (Bakker 1997; Muysken 1997).

5. References

Aboh, Enoch
2015 *The emergence of hybrid grammars.* Cambridge: Cambridge University Press.
Adger, David & Jennifer Smith
2005 Variation and the minimalist program. In Leonie Cornips & Karen P. Corrigan (eds.), *Syntax and variation: Reconciling the biological and the social*, 149–178. Amsterdam: John Benjamins.
Allen, Cynthia
1980 *Topics in diachronic English syntax*. New York: Garland.
Bakker, Peter
1997 *A language of our own: The genesis of Michif, the mixed Cree-French language of the Canadian Métis.* New York: Oxford University Press.
Bakker, Peter, Aymeric Daval-Markussen, Mikael Parkvall & Ingo Plag
2011 Creoles are typologically distinct from non-creoles. *Journal of Pidgin and Creole Languages* 26(1). 5–42.
Bickerton, Derek
1981 *Roots of language*. Ann Arbor, MI: Karoma Press.
Birdsong, David
1992 Ultimate attainment in second language acquisition. *Language* 68. 706–755.
Blasi, Damián E., Susanne Maria Michaelis & Martin Haspelmath
2017 Grammars are robustly transmitted even during the emergence of creole languages. *Nature Human Behaviour* 1. 723–729.
Bruhn de Garavito, Joyce & Lydia White
2000 L2 acquisition of Spanish DPs. In S. Catherine Howell, Sarah A. Fish & Thea Keith-Lucas (eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development*, 164–175. Somerville, MA: Cascadilla.
Camacho, José
2013 *Null subjects.* Cambridge: Cambridge University Press.
Chierchia, Gennaro
1998 Reference to kinds across languages. *Natural Language Semantics* 6. 339–405.
Chomsky, Noam
1982 *Some concepts and consequences of the theory of government and binding.* Cambridge, MA: MIT Press.
Cuza, Alejandro
2013 Crosslinguistic influence at the syntax proper. *International Journal of Bilingualism* 17. 71–96.

DeGraff, Michel
2003 Against creole exceptionalism. *Language* 79. 391–410.
de Granda, Germán
1968 La tipología criolla de dos hablas del área lingüística hispánica. *Thesaurus* 23. 193–205.
Epps, Patience & Lev Michael
2017 The areal linguistics of Amazonia. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 934–963. Cambridge: Cambridge University Press.
Franceschina, Florencia
2005 *Fossilized second language grammars*. Amsterdam: John Benjamins.
García Mayo, María del Pilar & Roger Hawkins
2009 *The second language acquisition of articles.* Amsterdam: John Benjamins.
Grimshaw, Jane & Vieri Samek-Lodovici
1998 Optimal subjects and subject universals. In Pilar Barbosa, Daniel Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the best good enough? Optimality and competition in syntax*, 193–219. Cambridge, MA: MIT Press.
Guerra Rivera, Alexia, Peter Coopmans & Sergio Baauw
2015 On the L2 acquisition of Spanish subject-verb inversion. *Procedia: Social and Behavioral Sciences* 173. 37–42.
Harris, Alice C. & Lyle Campbell
1995 *Historical syntax in cross-linguistic perspective.* Cambridge: Cambridge University Press.
Haugen, Einar
1950 The analysis of linguistic borrowing. *Language* 26. 210–231.
Holm, John
1992 Popular Brazilian Portuguese: A semi-creole. In Ernesto d'Andrade & Alain Kihm (eds.), *Actas do Colóquio sobre crioulos de Base Lexical Portuguesa*, 37–66. Lisboa: Edições Colibri.
Ionin, Tania & Silvina Montrul
2010 The role of L1-transfer in the interpretation of articles with definite plurals in L2-English. *Language Learning* 60(4). 877–925.
Jackendoff, Ray
1997 *The architecture of the language faculty.* Cambridge, MA: MIT Press.
King, Ruth
2000 *The lexical basis of grammatical borrowing*. Amsterdam: John Benjamins.
Kouwenberg, Silvia & John Victor Singler (eds.)
2008 *The handbook of pidgin and creole studies.* Oxford: Blackwell Publishing Ltd.
Lardiere, Donna
2007 *Ultimate attainment in second language acquisition: A case study.* Mahwah, NJ: Lawrence Erlbaum.
Lefevbre, Claire, Lydia White & Christine Jourdan (eds.)
2006 *L2 acquisition and creole genesis: Dialogues.* Amsterdam: John Benjamins.
Longobardi, Giuseppe
1994 Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25. 609–665.
Meillet, Antoine
1914 Le problème de la parenté des langues. *Scientia* 15. 403–425.
McWhorter, John
1998 Identifying the creole prototype. *Language* 74(4). 788–818.
McWhorter, John
2018 *The creole debate.* Cambridge: Cambridge University Press.
Montrul, Silvina
2004 *The acquisition of Spanish.* Amsterdam: John Benjamins.

Montrul, Silvina & Tania Ionin
2010 Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism, Language and Cognition* 13(4). 449–473.
Mufwene, Salikoko
1997 Jargons, pidgins, creoles, and koines: What are they? In Arthur Spears & Donald Winford (eds.), *The structure and status of pidgins and creoles*, 35–70. Amsterdam: John Benjamins.
Mufwene, Salikoko
2001 *The ecology of language evolution.* Cambridge: Cambridge University Press.
Muysken, Pieter
1981 Quechua en Spaans in het Andesgebied. *Tijdschrift voor Taal-en Tekstwetenschap* 1. 124–138.
Muysken, Pieter
1996 Syntax. In Hans Goebl, Peter H. Nelde, Zdeněk Starý & Wolfgang Wölck (eds.), *Contact linguistics: An international handbook of contemporary research*, vol. 1, 117–124. Berlin: Mouton de Gruyter.
Muysken, Pieter
1997 Media lengua. In Sarah G. Thomason (ed.), *Contact languages: A wider perspective*, 427–447. Amsterdam: John Benjamins.
Myers-Scotton, Carol
2002 *Contact linguistics. Bilingual encounters and grammatical outcomes.* Oxford: Oxford University Press.
Otheguy, Ricardo
1973 The Spanish Caribbean: A creole perspective. In Charles-James N. Bailey & Roger W. Shuy (eds.), *New ways of analyzing variation in English*, 323–339. Washington, DC: Georgetown University Press.
Picallo, M. Carme
1991 Nominals and nominalizations in Catalan. *Probus* 3. 279–316.
Pienemann, Manfred
1998 *Language processing and second language development: Processability theory.* Amsterdam: John Benjamins.
Pollock, Jean-Yves
1989 Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20. 365–424.
Rizzi, Luigi
1982 *Issues in Italian syntax.* Dordrecht: Foris.
Roberts, Ian
2007 *Diachronic syntax*. Oxford: Oxford University Press.
Rothman, Jason & Roumyana Slabakova (eds.)
2011 *Acquisition at the linguistic interfaces* [Special issue]. *Lingua* 121(4).
Sánchez Liliana & María José Giménez
1998 The L2 acquisition of definite determiners. In Annabel Greenhill, Mary Hughes, Heather Littlefield & Hugh Walsh (eds.), *Proceedings of the 22nd Annual Boston University Conference on Language Development*, 640–650. Somerville, MA: Cascadilla Press.
Schneider, Edgar
1990 The cline of creoleness in English-oriented creoles and semi-creoles of the Caribbean. *English World-Wide* 11(1). 79–113.
Sessarego, Sandro
2013a Afro-Hispanic contact varieties as advanced second languages. *Iberia: An International Journal of Theoretical Linguistics* 5(1). 96–122.
Sessarego, Sandro
2013b *Chota Valley Spanish.* Madrid/Frankfurt: Iberoamericana/Vervuert.

Sessarego, Sandro
2013c Enhancing dialogue between quantitative sociolinguistics and minimalist syntax. *Studies in Hispanic and Lusophone Linguistics* 5(2). 79–97.
Sessarego, Sandro
2014 *The Afro-Bolivian Spanish determiner phrase.* Columbus, OH: The Ohio State University Press.
Sessarego, Sandro
2015 *Afro-Peruvian Spanish.* Amsterdam: John Benjamins.
Sessarego, Sandro & Javier Gutiérrez-Rexach
2017 Revisiting the null subject parameter. *Isogloss* 3(1). 43–68.
Seuren, Pieter A.M. & Herman Wekker
1986 Semantic transparency as a factor in creole genesis. In Pieter Muysken & Norval Smith (eds.), *Substrata versus universals in creole genesis*, 57–70. Amsterdam: John Benjamins.
Slabakova, Roumyana
2008 *Meaning in the second language.* Berlin: Mouton de Gruyter.
Slabakova, Roumyana
2009 What is easy and what is hard to acquire in a second language? In Melissa Bowles, Tania Ionin, Silvina Montrul & Annie Tremblay (eds.), *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference*, 280–294. Somerville, MA: Cascadilla Press.
Sorace, Antonella
2004 Native language attrition and developmental instability at the syntax-discourse interface. *Bilingualism: Language and Cognition* 7. 143–145.
Sorace, Antonella & Ludovica Serratrice
2009 Internal and external interfaces in bilingual language development. *International Journal of Bilingualism* 13(2). 195–210.
Thomason, Sarah Grey & Terrence Kaufman
1988 *Language contact, creolization, and genetic linguistics.* Berkeley: University of California Press.
Tsimpli, Ianthi, Antonella Sorace, Caroline Heycock & Francesca Filiaci
2004 First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism* 8(3). 257–277.
van Coetsem, Frans
1988 *Loan phonology and the two transfer types in language contact.* Dordrecht: Foris.
Velupillai, Viveka
2015 *Pidgins, creoles and mixed languages.* Amsterdam: John Benjamins.
Weinreich, Uriel
1953 *Languages in contact.* Berlin: Mouton de Gruyter.
White, Lydia
2011 Second language acquisition at the interfaces. *Lingua* 121(4). 577–590.
White, Lydia & Alan Juffs
1998 Constraints on wh-movement in two different contexts of non-native language acquisition. In Suzanne Flynn, Gita Martohardjono & Wayne O'Neil (eds.), *The generative study of second language acquisition*, 111–129. Mahwah, NJ: Lawrence Erlbaum.
White, Lydia, Elena Valenzuela, Martyna Kozlowska-Macgregor & Yan-Kit Ingrid Leung
2004 Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics* 25(1). 105–133.
Whitney, William Dwight
1881 On mixture in language. *Transactions of the American Philosophical Association* 12. 1–26.

Winford, Donald
2000 "Intermediate" creoles and degrees of change in creole formation: The case of Bajan. In Ingrid Neumann Holzschuh & Edgar W. Schneider (eds.), *Degrees of restructuring in creole languages*, 215–246. Amsterdam: John Benjamins.
Winford, Donald
2003 *An introduction to contact linguistics*. Hoboken, NJ: Wiley-Blackwell.

Sandro Sessarego, Austin, TX (USA)

4. Contact-induced semantic change

1. Introduction
2. Fundamentals of semantic change
3. Semantic outcomes of contact: towards a typology
4. Dynamics of contact-induced semantic change
5. Semantic convergence and areal linguistics
6. Conclusion
7. References

1. Introduction

Contact-induced semantic change has received limited direct attention in the literature, in comparison to the lexical, morphosyntactic, and phonological effects of language contact. However, most outcomes of contact involve some change in how meaning is lexicalized or how elements combine to create meaning at the phrase and sentence levels in a language, and examples of such semantic effects are plentiful. The relationship between meaning and form in language contact is complex: form may be borrowed with or without the associated meaning from the donor language, and various aspects of meaning may be copied without the direct borrowing of form. Contact-driven changes in semantics may involve lexical elements, phrases, or whole functional domains through the development of new grammatical constructions and categories. The degree to which linguistic systems converge on the levels of meaning, form, and structure is sensitive to a wide range of factors, spanning cognitive processing, the maintenance of socially relevant distinctions, and discursive and cultural practices. Below, we introduce perspectives on meaning and diachrony, and survey the range of effects that contact across languages and language varieties may have on the organization and expression of meaning. We then focus in more detail on the role of cognitive and social variables in shaping these outcomes, and provide a brief discussion of the relationship between semantic change and linguistic areality.

https://doi.org/10.1515/9783110435351-004

2. Fundamentals of semantic change

Semantic representations have been understood according to at least two principal perspectives (e.g. Ameka and Wilkins 1996). Views that prioritize a 'dictionary' view of meaning consider words to correspond to a basic denotation, stored in the mind as purely linguistic knowledge; linguistic categories and structures are understood to be distinct from more general aspects of 'world' knowledge, and as "more or less straightforward mappings into our biological nature" (Li and Gleitman 2002: 226). In contrast, 'encyclopedic' views of meaning assume that any distinction between 'basic' linguistic knowledge and real-world knowledge is essentially arbitrary; meaning is dynamic and crucially dependent on context, with no strict divide between linguistic knowledge and language use. Approaches to meaning within this tradition make reference to 'frames' and 'domains', or schematizations of experience that may be associated with a particular word or other linguistic construction (e.g. Langacker 1987), and categories of meaning are associated with prototypes and linked to larger conceptual ontologies (as opposed to an Aristotelian structure based on necessary and sufficient components; see Lakoff 1986; Urban 2015). This perspective is implicit in such classic studies of semantic change as Spitzer's (1947) study of meaning in European languages and Matisoff's (1978) exploration of body-part terms in Sino-Tibetan.

'Encyclopedic' views of meaning are particularly attentive to cross-linguistic differences in semantic categories, how they may be structured and restructured over time and through contact with speakers of other languages, and the sensitivity of semantic change to cultural context (see Evans 2011: 505). This variability raises the problem of incommensurability in cross-linguistic studies of meaning, and highlights the challenges of translation (e.g. Evans and Sasse 2004). It also underscores how culturally specific categories of meaning might emerge and be shared across languages in situations of contact, as encoded in terms like 'one who has the evil eye' in languages of the Ethiopian region (Hayward 1991: 147).

Explorations into the diachrony of meaning are informed by the Comparative Method (e.g. Blust 1987; Urban 2015), and by a textual record where one exists (e.g. Traugott and Dasher 2002). Such investigations allow for the identification of cross-linguistically consistent patterns of change as well as more locally relevant ones (Evans and Wilkins 2000; Matisoff 1978; Wilkins 1996). An understanding of the processes and patterns behind semantic change is informed by links between diachrony and aspects of synchronic semantic structure, most notably with respect to polysemy (multiple senses), heterosemy (distinct senses associated with differences in word class or morphosyntactic slot), and degrees of vagueness or specificity (e.g. Evans 2011; Fortson 2003; Vanhove 2008). According to standard assumptions, the transition from one meaning to another normally involves a stage of polysemy in which both senses co-exist within the same form (e.g. Koptjevskaja-Tamm 2016; Sweetser 1990; Traugott and Dasher 2002; Urban 2015). New senses themselves emerge through the "semanticization of pragmatics" (Traugott 2010: 32) − that is, the process by which context-bound pragmatic inferences are reinterpreted as entailed meanings (see also Enfield 2003: 29), often involving lexical competition and markedness shift among polysemous terms (François 2011). In some cases, the emergence of 'bridging contexts' that enable these extensions may be facilitated by cultural convergence, while the reinterpretations themselves may take place independently (Evans 2003). The exploration of semantic change within the lexicon is still in the relatively

early stages (though see List et al. 2014; Zalizniak et al. 2012), but significant headway has been made in understanding the relationship between semantic and morphosyntactic change in processes of grammaticalization (Eckardt 2006; Heine and Kuteva 2005; Hopper and Traugott 2003). Assessing the role of language contact in processes of change offers a further set of methodological challenges: in addition to establishing that a change has in fact occurred, it must be demonstrated that contact with an appropriate model language or languages was a plausible driver (e.g. Campbell, Kaufman, and Smith-Stark 1986: 533–535).

3. Semantic outcomes of contact: towards a typology

The semantic outcomes of contact may be classified according to the ways that form and meaning in the donor language compare to form and meaning of the corresponding elements in the recipient language. A typology of possible outcomes must include both the borrowing of semantic structure without accompanying phonological form, and the borrowing of form together with full, partial, or no semantic structure (see Urban 2015: 377).

The borrowing of form and meaning together is generally treated as lexico-morphological rather than semantic borrowing, and is often presented as the default scenario. In reality, however, words are rarely borrowed with *exactly* the same referential range, polysemy patterns, and entailments of the donor language form. Language contact can also lead to the restructuring of lexical and/or grammatical meanings in the recipient language, independent of the borrowing of forms (see also Chapters 13 and 14). Weinreich ([1953] 1966: 1) terms such processes ‘interference’, in contrast to ‘borrowing’ of form and meaning together, while more recent work by Matras and Sakel (2007) refers to ‘MAT (matter)’ vs. ‘PAT (pattern)’ borrowing. The diffusion of semantic and grammatical structure is often associated with linguistic areality (see Section 5). We consider instances of borrowing of meaning that are relatively independent of form, followed by a look at semantics in the context of the borrowing of form *and* meaning.

3.1. Borrowing of meaning without form

Much semantic borrowing is not accompanied by the borrowing of phonological forms, although much of what has been said explicitly about contact semantics focuses on cases where form is also involved. The restructuring of meaning can occur at the level of specific words and set phrases, or in the organization of grammatical categories, semantic fields, and discourse.

At the lexical level, a language can copy additional senses of individual morphemes, leading to converging patterns of polysemy (Matisoff 2004; Urban 2015; Vanhove 2008; Zalizniak et al. 2012). For example, Enfield (2003) gives a detailed description of the areally shared polysemy patterns of words meaning ‘acquire’ in Southeast Asian language, which have developed further meanings relating to possibility (Example 1), achievement, aspect marking, etc. Multimorphemic expressions (whether words or phrases) can also be translated morpheme-by-morpheme, as seen for example in the pan-

Mesoamerican term for ‘wrist’ as ‘neck of the arm’ (Campbell, Kaufman, and Smith-Stark 1986). Such processes of semantic extension and calquing may be mediated by syntactic parallels across the languages, and over time may facilitate the transfer of structural features.

(1) a. Lao
hêt[1] *naa*[2] *bò*[0] ***daj***[4]
do/make paddy.rice.field NEG can
‘one can’t work rice paddies (there)’

b. Cantonese
jáu ***dāk*** *ge* *lak*
leave can PARTICLE PARTICLE
‘(we) can leave now’

c. Modern Standard Chinese
yào *bu* ***dé***
want NEG can
‘cannot be wanted, undesirable’
(Enfield 2003: 196–197)

Contact may also lead to the semantic reorganization of grammatical categories. Such changes can be seen in the noun classification systems of the Amazonian Vaupés region; for instance, Baniwa (Arawakan) influence has caused Cubeo (East Tukanoan) to change its strategies for classifying animate entities from prioritizing animacy (the Tukanoan pattern) to shape (the Arawakan pattern; Gomez-Imbert 1996), while Tariana (Arawakan) has experienced exactly the opposite change under the influence of Tukano (East Tukanoan), as can be seen by comparison with Tariana’s close relative Baniwa (Example 2).

(2) a. Tariana
api *hanu-ite*
snake big-CLF:ANIMATE
‘a big snake’

b. Tukano
pĩrõ *pahi-gɨ́*
snake big-CLF:ANIMATE
‘a big snake’

c. Baniwa
a:pi *maka-khay*
snake big-CLF:CURVILINEAR
‘a big snake’
(Aikhenvald 2002: 89–92)

When the contribution of contact is a novel meaning from a combination of native words, or extends or redistributes polysemy patterns within a portion of the lexicon or within a grammatical domain, we can consider this effect to involve *conceptual reorganization*. However, many instances of lexical borrowing in particular are motivated by the intro-

duction of entirely new fields of knowledge, experience, and discourse (local flora/fauna, computers, generative linguistics). Such cases may involve adding new concepts rather than simply reshuffling the lexical organization of an existing conceptual field. This *conceptual expansion* can be accomplished in many different ways. In some cases, new words are borrowed or coined. In others, old words are put to new uses, yielding polysemy or a new meaning through a process of markedness shift. Either way, such processes of semantic change typically involve inter-cultural communication and exchange.

3.2. Borrowing of form with(out) meaning

The borrowing of lexical and morphological material also involves semantics, in that it typically entails the copying of at least some part of the semantic and grammatical properties of the borrowed element. Even in cases where words are borrowed as proper names without the associated lexical meaning found in the donor language (e.g. *Daniel*, *Mississippi*), at least one referential element of the source language meaning is retained: that of proper naming. In some cases, loanwords are borrowed with meanings that closely match the referential range of the word in the donor language, while in others the meaning is significantly changed. For example, the Spanish word *queso* 'cheese' was borrowed into the English of the Southwestern United States, where it does not refer to cheese generally, but to a specific type of cheese dip flavored with chili peppers.

The differences that we find between recipient and donor language semantics for loanwords are as diverse as the range of shifts that have been described and documented in language-internal semantic change, including semantic widening, narrowing, metaphorical extensions, and relations of metonymy or synecdoche. However, more empirical work is needed to determine whether certain patterns of semantic change are favored in contact conditions, and whether they are likely to occur during and/or after a process of borrowing and nativization. We might expect a loanword to have a more restricted range of reference than its source word in the donor language (Mańczak 1985: 367). In practice, however, it is quite difficult to quantify the number of 'meanings' or senses that a word has. We must also distinguish the number of distinct meanings (polysemy) from the degree to which the meaning may simply be vague, and/or one meaning may contain another (e.g. via processes of narrowing or widening; Geeraerts 1993; Riemer 2005). While borrowed words may be more specific in the recipient language than in the donor language (as with *queso*), they may also be more general (often through subsequent widening), as in Bulgarian *penkiler* 'something/someone that can do anything, jack of all trades', from English *painkiller* (Alexieva 2008).

Semantics also plays a role in determining the susceptibility of a morphological or lexical element to borrowing. Recent work (e.g. Haspelmath and Tadmor 2009; Wohlgemuth 2009) has empirically strengthened and refined the long-held hypothesis that lexical borrowing occurs at different rates across different word classes and semantic domains, particularly those involving 'basic' vocabulary versus other categories, such as flora-fauna and cultural terms, and more versus less referential elements, such as nouns or verbs.

In addition to the effects of semantic contact on the lexicon, languages may develop or restructure meanings together with new grammatical categories and functions, often in conjunction with syntactic restructuring. Examples include the reorganization of tense

values in a language, as in the development in Hup (Naduhupan) of a recent/remote past distinction from what appears to have been a single generic past tense category (Epps 2007), and the introduction of new grammatical categories, such as numeral classifiers in Lowland Mayan languages (Law 2014; Example 3).

(3) Tseltal
*Ox-**wol*** *y-u'un* *ox-**jil*** *pak'* *x-lok'* *ts'in* *abi.*
three-**CLF:BALL** 3ERG-RN three-**CLF:CLOTH** cloth IPFV-go.out then EVID
'(There are) three (balls of thread) so that three lengths of cloth will come out of it.'
(Polian 2013: 591; our translation from Spanish)

4. Dynamics of contact-induced semantic change

We are still far from a robust understanding of the mechanisms by which language contact may propel semantic change. The range of outcomes addressed above indicates that their motivations are variable and sensitive to many factors. Here we review aspects of the processes and pathways by which contact-driven semantic change may take place, and some of the cognitive, social, and other forces behind those changes.

4.1. Processes and pathways

Semantic change involves the development of polysemy, as implicatures become routinized and give rise to new entailed meanings (e.g. Deo 2015; Koptjevskaja-Tamm 2016; Urban 2015); however, contact provides particular mechanisms that direct and even catalyze these processes. Speakers in multilingual contexts are already primed for certain entailments, whether cued partly from pragmatic context or building on a shared awareness of the model in the contact language. The presence of the model in the source language is therefore likely to both promote particular extensions of meaning in the recipient language, and make the move from implicature to entailment rapid for individual speakers and across a speech community. The strength of these effects is linked to the pervasiveness of multilingualism across the speech community.

One process of contact-driven semantic change is described as 'polysemy calquing' or 'copying'. For example, Rumsey (1983: 183) describes conflations of meanings in north Australian Kriol that are motivated by semantic links in the traditional languages, such as *jigrid* < 'sacred/secret'. Polysemy copying is frequently invoked to account for congruence in semantic categories across languages in contact, such as that of 'eat' = 'drink' in many Papuan and Australian Aboriginal languages (Koptjevskaja-Tamm and Liljegren 2017), and 'mouse' = 'computer accessory' in various European languages.

Polysemy copying may also come into play in contact-induced grammaticalization, where structure and semantics are modified in tandem (Enfield 2003; Heine and Kuteva 2005). However, as Heine and Kuteva (2005) observe, polysemy copying alone cannot explain why features of a category that emerges in the recipient language are often not identical to that of the source language. Examples of such mismatches in contact-induced developments are common, particularly in those involving grammaticalization, where

semantic change accompanies morphosyntactic restructuring. For example, the emergence of a set of evidential categories in the Vaupés language Hup has clearly been motivated by contact with Tukanoan, but despite a very close match in this domain between the donor and the recipient languages, there are subtle differences, such as Hup speakers' preferred use of the inferred evidential to mark accidental, uncontrollable actions, whereas Tukano speakers tend to use the non-visual form for such actions (Epps 2005; see also Aikhenvald 2002: 119).

Such discrepancies between donor and recipient languages are expected, given that contact-induced change involves a combination of variably-applied external and internal processes. Heine and Kuteva's (2005) discussion of contact-induced grammaticalization offers several scenarios by which these processes may co-occur. In some cases, they argue, speakers may observe a grammatical category in the model language, and then develop their own version via language-internal (and/or universally relevant) strategies. In others, they suggest that speakers may consciously replicate a grammaticalization sequence that they assume to have taken place in the other language. However, such a teleological explanation has been questioned (Heine 2012; Ross 2007). According to Ross' (2007: 135) formulation, multilingual speakers are more likely to "establish functional correspondences between the constructions of the two languages, and remodel the constructions of language R so that they more closely match those of language M […]. These adaptations may in turn lead to grammaticisation". Regardless of the precise dynamics involved, such complex interactions between (relatively) externally and internally motivated processes are clearly relevant to changes in meaning, as well as in structure and form.

Contact-driven changes in meaning may also be sensitive to elements of form, even where form itself is not borrowed. This phenomenon is evident in cases of lexical-semantic 'contamination'; for example, the meaning of Old English *drēam* changed from 'joy' to 'vision while asleep' under the influence of Scandinavian *draumr*, which had the latter meaning while a cognate with the English form (Algeo 2010: 253). In a more complex case, parallel constellations of meanings and grammatical function have developed around forms resembling *ni* in languages of at least three different families of the Amazonian Vaupés, ranging from a verb of existence/location, a copula in equative clauses, an evidential, an aspectual marker, a light verb in a predicate chaining construction, and a verbalizing derivational element (Epps 2012). Though the *ni* etyma appear to have varied origins in these languages, the diffusion of grammatical and semantic features has clearly been facilitated by their formal similarity, despite the fact that speakers in this region tend to resist the direct borrowing of form.

4.2. Motivations in cognition and interaction

The motivations for contact-driven semantic change are as varied and complex as the processes involved. One important area of exploration is the cognitive or psychological dimension of the multilingual mind. Questions relating to how linguistically distinct forms and their associated concepts are stored have been key concerns in research into multilingualism, going back to Weinreich's (1966) proposed distinction between coordinate bilingualism (in which words in distinct languages are associated with distinct concepts) and compound (or subordinate) bilingualism (in which words in distinct languages

may be associated with a single concept). Subsequent research has made it clear that the organization of forms and concepts in the multilingual mind is vastly more complex than this two-way division suggests, and that many factors shape that complexity, including the setting or cultural context in which the languages are acquired, the age of acquisition, etc. (e.g. Ameka and Wilkins 1996; Ervin and Osgood 1954; Grosjean 2001; Muysken 2000). Recent work on contact between colonial and indigenous languages indicates that bilingualism may also be 'elaborative', such that the set of terms resulting from the contact situation realizes the full set of distinctions found in both languages (e.g. Dickson 2015).

The role of the cognitive load presented by the task of processing multiple languages has also received attention in studies of multilingualism and contact-driven convergence (e.g. Appel and Muysken 1987; Ross 2007). An extensive literature on second language acquisition highlights the phenomenon of interference or transfer, where structures of the first language consistently influence the acquisition of the second (Doughty and Long 2003; Odlin 1989), which can translate into pervasive community-wide effects in situations of substrate influence. Such cognitive processes are certainly part of the explanation for contact-driven convergence on multiple levels, including semantics, and also help to account for differential convergence across levels − for example, higher borrowing rates among utterance modifiers and other discourse-managing elements (Matras 2007), and situations where the borrowing of phonological forms is closely constrained while structural and semantic convergence is pervasive (Aikhenvald 2002; Ross 2007; Silverstein 1981). The cognitive load involved in keeping languages distinct may also contribute to *divergence*, as Ellison and Miceli (2017) show in their study of 'doppel avoidance' − the phenomenon by which bilingual speakers unconsciously *avoid* using etyma that are formally similar across the two languages. François (2011) explores how such forces favoring linguistic distinction − particularly at the level of phonological form − may work together with those favoring convergence, such that lexical innovation may be followed by semantic restructuring, bringing the new etymon into line with areal expectations regarding lexicalized categories of meaning. An example is the word meaning 'random' (Example 4) and its semantic extensions 'casually', 'ordinary', etc., found across languages of northern Vanuatu.

(4)

Motwap	*kɪ*	*ni-*	*ⁿdɛlⁿdɛl*	***sɔlɔtɛɣ***		*lɛ-*	*pnʊ*
Lemerig	*ti*	*n-*	*tɛltɛl*	***βʊlʊkrɛɣ***		*lɪ*	*βunu*
Koro	*ni*	*t-*	*ɛlɛ̃al*	***kɪlkɪl***	*ti*	*l-*	*βʊnʊ*
Lakon	*ni*	*tɪ*	*βanβan*	***sælɪn***	*tɔ*	*lɪ*	*βanʊ*
	3SG	IPFV	wander~RED	random	[$IPFV_2$]	LOC	village

'He was wandering *around* in the village'.

(François 2011: 215)

While cognitive processes are understood to be universally relevant, their role in structuring contact-driven semantic change is dependent on the cultural and social context in which the languages are used. The degree to which multilingual speakers strive to keep their languages distinct is linked to local linguistic ideologies, relating to whether and to what extent mixing is noticed and tolerated. The striking combination of structural/semantic convergence with maintenance, and even divergence, of phonological form in regions like Vanuatu and the Amazonian Vaupés are linked to particular ideological

alignments of social and linguistic identity (such that unintentional or otherwise non-conventional language mixing is stigmatized), whereas ideological motivations are argued to have facilitated language mixing in the Maya Lowlands (Law 2014).

At the same time, the degree of multilingualism necessary for extensive convergence to occur is grounded in the social and historical context of interaction among speakers over time. This interaction involves discursive practices that foster particular ways of speaking. As Enfield (2011: 36) observes in his discussion of convergence in the lexical-semantic domain of taste in Lao and Kri (southeast Asia), "when a language has a word for a specific thing, it is not because of the presence of this thing in the environment, but because speakers have historically engaged in social practices of using that word [… What is] transmitted is a practice of using the word, and a recurring set of contexts for such use." The relevance of discursive practice is also evident in calquing. In Vaupés languages, in fact, it is possible to associate particular calqued formations to specific discursive genres: for example, the convergence of tense-evidential marking appears to be associated with traditional narrative (Epps 2005).

5. Semantic convergence and areal linguistics

Evidence of contact-driven convergence of semantic categories is often widespread within a region, affecting multiple languages. Congruence in semantic categories is discussed as a diagnostic for linguistic areality in a number of regions, such as the Balkans (Friedman 2007), Ethiopia (Hayward 1991; Treis 2010), Mesoamerica (Campbell, Kaufman, and Smith-Stark 1986), southeast Asia (Matisoff 2004), and the Circum-Baltic region (Dahl and Koptjevskaja-Tamm 2001); see also Koptjevskaja-Tamm and Liljegren (2017). Examples of convergent meanings in these regions include: calques, such as forms meaning 'jaguar-cacao' (for 'wild cacao') and 'mother of driver ant' (for 'coral snake') in a number of Mesoamerican languages; shared semantic specializations, like 'die without ritual slaughter (of cattle)' in Ethiopian languages (Hayward 1991); and congruent polysemies, like 'deer' = 'manioc-washing tripod' in languages of the Amazonian Vaupés. Semantic similarities, like other evidence of contact, may be tightly localized or much more widely distributed within a given region where contact is understood to have taken place, and may be indicative of different contexts and levels of interaction (e.g. Friedman 2007). In the Amazon basin, for example, languages within localized contact zones like the Vaupés share a large number of calques and other semantic loans, but at least a few examples of such forms are widespread across much of lowland South America, such as a term for 'four' built on words for 'companion' or 'sibling' (Epps 2013; Example 5).

(5)	Hup (Naduhupan)	*hibab'ni*	'have sibling, be accompanied, 4'
	Tuyuka (East Tukanoan)	*bapari*	'companions, friends, 4'
	Miraña (Boran)	*tsanɨ́ná?bɛ́ɛ̨ahkátsí*	'being companions to each other, 4'
	Iquito (Zaparoan)	*suhuaramaajɨtáami*	'having proper same-sex sibling, 4'

In extreme cases, semantic convergence and heavy morphosyntactic restructuring may work together to yield morpheme-by-morpheme equivalence and inter-translatability be-

tween two or more languages, while distinctions in phonological form are maintained or even exaggerated (see Thomason 2007). A well-known example of such cross-linguistic isomorphism is Gumperz and Wilson's (1971) discussion of Kupwar village in India. Extreme convergence of this kind has been labeled 'metatypy' by Ross (2007) in his investigation of contact-driven restructuring in Takia (Oceanic) and Waskia (Papuan); see Example (6), where these are contrasted to Takia's Oceanic sister Arop-Lokep. Pervasive structural and semantic convergence has also been observed among unrelated languages of the Amazonian Vaupés region, coupled with low levels of lexical borrowing (Aikhenvald 2002; Epps 2007; Gomez-Imbert 1996). François's (2011) discussion of structural isomorphism among 17 related Oceanic languages of Vanuatu likewise notes their high divergence in phonological form. Such cases challenge earlier assumptions that grammatical restructuring could only occur in the context of heavy lexical borrowing (e.g. Sankoff 2002).

(6) a. Takia (Oceanic)
Kai sa-n ab
Kai CLF-his house
'Kai's house'

b. Waskia (Papuan)
Kai ko kawam
Kai ABL house
'Kai's house'

c. Arop-Lokep (Oceanic)
rumu ke tool in
house ABL man that
'that man's house'
(Ross 2007: 118–119)

Particularly from an 'encyclopedic' view of semantics, we must confront the problematic boundary between semantic contact and cultural contact, or simply the sharing of knowledge more generally. For example, Mesoamerican indigenous languages share a calendar of 20 days, in which many of the day names are derived from names for animals (e.g. 'deer' and 'snake') in the various languages that use the calendar. While we can view these calqued day names as examples of semantic borrowing relating to calendrical values, what has been shared goes beyond the names of the days themselves, and includes the *concept* of naming days and a particular calendrical system. As new philosophies and technologies are shared around the world, new domains of knowledge are shared along with new lexical items. One challenge for the future will be to clarify empirically the relationship between semantic contact and the diffusion of knowledge and cultural practices more generally.

6. Conclusion

Considerable progress has been made in exploring the processes, patterns, and motivations for contact-driven semantic change over the past few decades. We now have signifi-

cantly more breadth and depth in synchronic descriptions of the world's languages, which in turn has fed more fine-grained investigations of their histories. This work has produced essential insights regarding the extent to which rates and processes of semantic change are recurrent and/or variable across regions and language families. Nevertheless, we are still a long way from reaching a robust understanding of semantic change in contact conditions.

The field of contact semantics offers many promising directions for future research, and work in a number of these areas is already well underway. The ongoing development of extensive databases cataloguing semantic shifts and co-lexifications in the world's languages (List et al. 2014; Zalizniak et al. 2012) will provide rich material for studies of diachronic semantics and the influences of contact on these processes. Enfield (2011) points out the need for more in-depth investigation into differences in individuals' internal representations of word meaning, and of the importance of a focus on language use in everyday interactions in exploring these questions. Recent investigations of the cognitive processes behind both convergence and divergence likewise suggest many new angles to pursue, as does ethnographically engaged work on the social and cultural context of language contact. These and many other half-open doors offer intriguing glimpses into future discoveries in semantic contact.

7. References

Aikhenvald, Alexandra Y.
2002 *Language contact in Amazonia*. Oxford & New York: Oxford University Press.
Aikhenvald, Alexandra & R. M. W. Dixon (eds.)
2007 *Grammars in contact: A cross-linguistic typology*. Oxford: Oxford University Press.
Alexieva, Nevena
2008 How and why are Anglicisms often lexically different from their English etymons? In Rowena Fischer & Hanna Pułaczewska (eds.), *Anglicisms in Europe: Linguistic diversity in a global context*, 42–51. Newcastle Upon Tyne: Cambridge Scholars.
Algeo, John
2010 *The origin and development of the English language*, 6th edn. Boston, MA: Wadsworth, Cengage Learning.
Ameka, Felix & David Wilkins
1996 Semantics. In Hans Goebl, Peter H. Nelde, Zdenek Starý & Wolfgang Wölck (eds.), *Contact linguistics*, Vol. 1, 130–138. Berlin & New York: Walter de Gruyter.
Appel, Rene & Pieter Muysken
1987 *Language contact and the bilingual mind*. London: Edward Arnold.
Blust, Robert
1987 Lexical reconstruction and semantic reconstruction: The case of Austronesian 'house' words. *Diachronica* 4. 79–106.
Campbell, Lyle, Terrence Kaufmann & Thomas Smith-Stark
1986 Mesoamerica as a linguistic area. *Language* 62. 530–570.
Dahl, Östen & Maria Koptjevskaja-Tamm (eds.)
2001 *The Circum-Baltic languages: Typology and contact*, 2 vols. Amsterdam: John Benjamins.
Deo, Ashwini
2015 Diachronic semantics. *Annual Review of Linguistics* 1. 179–197.

Dickson, Gregory
2015 Marra and Kriol: the loss and maintenance of knowledge across a language shift boundary. PhD dissertation, Australian National University.
Doughty, Catherine J. & Michael H. Long (eds.)
2003 *The handbook of second language acquisition*. Oxford: Blackwell.
Eckardt, Regine
2006 *Meaning change in grammaticalization: An enquiry into semantic reanalysis*. Oxford: Oxford University Press.
Ellison, T. Mark & Luisa Miceli
2017 Contact-induced lexical differentiation: Bilingual monitoring leads to a loss of shared lexical forms. *Language* 93(2). 255–287.
Enfield, Nicholas J.
2003 *Linguistic epidemiology*. Cambridge: Cambridge University Press.
Enfield, Nicholas J.
2011 Taste in two tongues: A Southeast Asian study of semantic convergence. *The Senses and Society* 6(1). 30–37.
Epps, Patience
2005 Areal diffusion and the development of evidentiality: Evidence from Hup. *Studies in Language* 29(3). 617–650.
Epps, Patience
2007 The Vaupés melting pot: Tucanoan influence on Hup. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Grammars in contact: A cross-linguistic typology*, 267–289. Oxford: Oxford University Press.
Epps, Patience
2012 On form and function in language contact: A case study from the Amazonian Vaupés region. In Isabelle Léglise & Claudine Chamoreau (eds.), *Dynamics of contact-induced language change*, 195–230. Berlin: de Gruyter Mouton.
Epps, Patience
2013 Inheritance, calquing, or independent innovation? Reconstructing morphological complexity in Amazonian numerals. *Journal of Language Contact* 6. 329–357.
Ervin, Susan & Charles E. Osgood
1954 Second language learning and bilingualism. In Charles E. Osgood & Thomas A. Sebeok (eds.), *Psycholinguistics: a survey of theory and research problems*, 139–146. Bloomington, Indiana: Indiana University Press.
Evans, Nicholas
2003 Context, culture and structuration in the languages of Australia. *Annual Review of Anthropology* 32. 13–40.
Evans, Nicholas
2011 Semantic typology. In Jae Jung Song (ed.), *The Oxford handbook of typology*, 504–553. Oxford: Oxford University Press.
Evans, Nicholas & David Wilkins
2000 In the mind's ear: Semantic extensions of perception verbs in Australian languages. *Language* 76(3). 546–592.
Evans, Nicholas & Hans-Jürgen Sasse
2004 Searching for meaning in the library of Babel: Field semantics and problems of digital archiving. In Linda Barwick & Jane Simpson (eds.), *Proceedings of the Paradisec Workshop on Digital Archiving*. Online at http://www.paradisec.org.au/EvansSasse_paper_rev1.pdf. (accessed on 15 August 2018).
Fortson, Benjamin. W., IV
2003 An approach to semantic change. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 648–666. Oxford, UK: Blackwell Publishing Ltd.

François, Alexandre
2011 Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics* 1(2). 175–246.
Friedman, Victor
2007 Balkanizing the Balkan Sprachbund: A closer look at grammatical permeability and feature distribution. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Grammars in contact: A cross-linguistic typology*, 205–220. Oxford: Oxford University Press.
Geeraerts, Dirk
1993 Vagueness's puzzles, polysemy's vagaries. *Cognitive Linguistics* 4(3). 223–272.
Gomez-Imbert, Elsa
1996 When animals become 'rounded' and 'feminine': Conceptual categories and linguistic classification in a multilingual setting. In John Gumperz & Stephen Levinson (eds.), *Rethinking linguistic relativity*, 438–469. Cambridge: Cambridge University Press.
Grosjean, Francois
2001 The bilingual's language modes. In Janet L. Nicol (ed.), *One mind, two languages: Bilingual language processing*, 1–22. Oxford: Blackwell.
Gumperz, John J. & Robert Wilson
1971 Convergence and creolization: A case from the Indo-Aryan/Dravidian border. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 151–167. London: Cambridge University Press.
Haspelmath, Martin & Uri Tadmor (eds.)
2009 *Loanwords in the world's languages: A comparative handbook.* Berlin: Mouton de Gruyter.
Hayward, Richard
1991 A propos patterns of lexicalization in the Ethiopian language area. In Daniela Mendel & Ulrike Claudi (eds.), *Ägypten im afroorientalischen Kontext* (Special issue of *Afrikanistische Arbeitspapiere*), 139–156. Cologne: Institute of African Studies.
Heine, Bernd
2012 On polysemy copying and grammaticalization in language contact. In Claudine Chamoreau & Isabelle Léglise (eds.), *Dynamics of contact-induced language change*, 125–166. Berlin: De Gruyter Mouton.
Heine, Bernd & Tania Kuteva
2005 *Language contact and grammatical change*. Cambridge: Cambridge University Press.
Hopper, Paul J. & Elizabeth C. Traugott
2003 *Grammaticalization*, 2nd edn. Cambridge: Cambridge University Press.
Koptjevskaja-Tamm, Maria
2016 The lexical typology of semantic shifts: An introduction. In Päivi Juvonen & Maria Koptjevskaja-Tamm (eds.), *The lexical typology of semantic shifts*, 1–20. Berlin: de Gruyter Mouton.
Koptjevskaja-Tamm, Maria & Henrik Liljegren
2017 Semantic patterns from an areal perspective. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 204–263. Oxford: Oxford University Press.
Lakoff, George
1986 Classifiers as a reflection of mind. In Colette G. Craig (ed.), *Noun classes and categorization*, 13–51. Amsterdam: John Benjamins.
Langacker, Ronald W.
1987 *Foundations of cognitive grammar: Theoretical prerequisites.* Stanford, CA: Stanford University Press.
Law, Danny
2014 *Language contact, inherited similarity and social difference: The story of language contact in the Maya lowlands.* Amsterdam: John Benjamins.

Li, Peggy & Lila Gleitman
2002 Turning the tables: Language and spatial reasoning. *Cognition* 83. 265–294.
List, Johann-Mattis, Thomas Mayer, Anselm Terhalle & Matthias Urban
2014 CLICS: Database of cross-linguistic colexifications. Marburg: Forschungszentrum Deutscher Sprachatlas (Version 1.0, available at http://CLICS.lingpy.org).
Mańczak, Witold
1985 Semantic development of borrowings. In Jacek Fisiak (ed.), *Historical semantics, historical word-formation*, 367–375. Berlin: Mouton de Gruyter.
Matisoff, James
1978 *Variational semantics in Tibeto-Burman*. Philadelphia: Institute for the Study of Human Issues.
Matisoff, James
2004 Areal semantics: Is there such a thing? In Anju Saxena (ed.), *Himalayan languages, past and present*, 347–393. Berlin: de Gruyter.
Matras, Yaron
2007 The borrowability of structural categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 31–69. Berlin: Mouton de Gruyter.
Matras, Yaron & Jeannette Sakel
2007 Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31. 829–865.
Muysken, Pieter
2000 *A typology of code-mixing*. Cambridge: Cambridge University Press.
Odlin, Terence
1989 *Language transfer: Cross-linguistic influence in language learning.* Cambridge: Cambridge University Press.
Polian, Gilles
2013 *Gramática del tseltal de Oxchuc*. Mexico City: CIESAS.
Riemer, Nick
2005 *The semantics of polysemy: Reading meaning in English and Warlpiri*. Berlin & New York: Mouton de Gruyter.
Ross, Malcolm
2007 Calquing and metatypy. *Journal of Language Contact − THEMA* 1. 116–143.
Rumsey, Alan
1983 On some syntactico-semantic consequences of homophony in north-west Australian Pidgin/Creole English. *Papers in Pidgin and Creole Linguistics* 3. 177–189.
Sankoff, Gillian
2002 Linguistic outcomes of language contact. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *Handbook of language variation and change*, 338–368. Oxford: Blackwell.
Silverstein, Michael
1981 The limits of awareness. *Working Papers in Sociolinguistics* no. 84. Austin, TX: Southwest Educational Development Laboratory.
Spitzer, Leo
1947 *Essays in historical semantics*. New York: Russell & Russell.
Sweetser, Eve
1990 *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
Thomason, Sarah Grey
2007 Language contact and deliberate change. *Journal of Language Contact* 1. 41–62.

Traugott, Elizabeth C.
2010 (Inter)subjectivity and (inter)subjectification: A reassessment. In Kristin Davidse, Lieven Vandelanotte & Hubert Cuyckens (eds.), *Subjectification, intersubjectification and grammaticalization*, 29–70. Berlin: Walter de Gruyter.
Traugott, Elizabeth C. & Richard B. Dasher
2002 *Regularity in semantic change.* Cambridge: Cambridge University Press.
Treis, Yvonne
2010 Perception verbs and taste adjectives in Kambaata and beyond. In Anne Storch (ed.), *Perception of the invisible: Religion, historical semantics and the role of perceptive verbs* (SUGIA 21), 313–346. Cologne: Köppe.
Urban, Matthias
2015 Lexical semantic change and semantic reconstruction. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 374–392. New York: Routledge.
Vanhove, Martine (ed.)
2008 *From polysemy to semantic change*. Amsterdam: John Benjamins.
Viberg, Åke
1984 The verbs of perception: A typological study. In Brian Butterworth, Bernard Comrie & Östen Dahl (eds.), *Explanations for language universals*, 123–162. Berlin: Mouton de Gruyter.
Weinreich, Uriel
1966 [1953] *Languages in contact.* The Hague: Mouton.
Wohlgemuth, Jan
2009 *A typology of verbal borrowings*. Berlin & New York: Mouton de Gruyter.
Wilkins, David
1996 Natural tendencies in semantic change and the search for cognates. In Mark Durie & Malcolm Ross (eds.), *The Comparative Method reviewed*, 264–304. Oxford: Oxford University Press.
Zalizniak, Anna, Maria Bulakh, Dimiotrij Ganenkov, Ilya Gruntov, Timur Maisak & Maxim Russo
2012 The catalogue of semantic shifts as a database for lexical semantic typology. *Linguistics* 50(3). 633–669.

Patience Epps, Austin, TX (USA)
Danny Law, Austin, TX (USA)

5. Lexicon and word formation

1. Introduction
2. Contexts
3. Courses
4. Consequences
5. Conclusions
6. References

1. Introduction

This chapter looks at processes of vocabulary expansion in general, and word formation in particular, in languages under contact. Section 2 (Contexts) examines the social and

https://doi.org/10.1515/9783110435351-005

linguistic predictors of lexical change due to contact, and the motivations for vocabulary adoption and adaptation. Section 3 (Courses) examines particular processes of word formation, (ad)lexification, and lexical integration, taking into account productivity in language and constraints on it. Section 4 (Consequences) looks at the effects of contact on the lexicon, both synchronic (lexical layers and expressivity) and diachronic (vocabulary richness and speaker awareness). The last section (Conclusions) questions how recent developments in linguistics, including corpus-based and cognition studies, can inform our understanding of lexical creativity in general, and of word formation due to contact in particular.

2. Contexts

As the native vocabularies of the world's languages grow, they rarely remain self-sufficient, attracting instead elements or forms from other languages. And as language contact seems to be the norm, borrowing is a common means of vocabulary expansion. According to Geeraerts (2015: 422), "borrowing competes with morphological word formation for being the most common, the most productive lexicogenetic mechanism". In other words, lexical creativity is furnished mainly either by use of native morphology (and processes like derivational affixation) or by introduction of foreign loans, and less so by pure neologism or by transformation of existing words (through processes such as clipping, blending, and acronym creation).

Haugen (1950) and Fischer (2003) give detailed descriptions of the typology of loans. Loanwords, perhaps the most common type of vocabulary expansion through contact, are whole words borrowed from another language in form and meaning. *Yoghourt* from Turkish and *sushi* from Japanese are examples of loanwords in English. Loan translations, or calques, are replications of the structure of foreign language words or expressions using synonymous native word forms. *Uranoksístis*, in Greek, is a loan translation of the English 'skyscraper', a compound of *uranós* ('sky') and *ksíno* ('to scrape'). Loan blends combine both direct borrowing and native substitution of morphs in complex borrowed words. The German word *Grüngrocer* was coined by combining the German *grün* ('green'), with the English 'grocer'. Semantic loans are words whose meaning is extended as a result of association with the meaning of a partly synonymous word in a foreign language (see Chapter 4 for details on contact-induced semantic change). For example, the English verb 'to realize' means both 'to make something become true' and 'to become aware of something'. The German verb *realisieren* originally just had the first meaning, but then later borrowed the second meaning from English and attached it to the existing word. Durkin (2009) explains the difficulty of differentiating loan translations and semantic loans from coincidental parallels in two languages. Durkin (2006) and Miller (2014) discuss the difficulty of dating borrowings and neologisms, as well as some solutions to the problem.

Grant (2015) summarizes numerous studies of language contact and linguistic borrowing. He writes: "Prevailing thought holds that speakers of culturally prestigious or politically powerful languages ('donor languages') bring about changes upon the languages ('recipient languages') used by less powerful populations, with acquisition of labels for previously unfamiliar items preceding the replacement by borrowed labels of

pre-existing labels for familiar concepts" (Grant 2015: 431–432). This statement also repeats the two main motivations for borrowing, need and prestige, and the two main categories of loans, cultural and replacive (or Haspelmath's [2009] 'core') borrowing. As far as motivations are concerned, 'need' drives much lexical borrowing in situations of culture and language contact. 'Prestige', problematic because it is hard to define and tends to oversimplify complex sociolinguistic situations (Milroy 1989), is sometimes still used in contact studies. As far as categories are concerned, it is worth considering them with regard to their effect on the lexicon at large: cultural, or 'necessary', borrowing results in adlexification, vocabulary growth par excellence. Replacive, or 'unnecessary', borrowing may result in supralexification, a term coined by Hancock (1971) to capture the semantic complications of a native vocabulary post-contact (more on this later). The former involves onomasiological change, or addition of labels to an existing lexicon. The latter involves semasiological change, or addition or modification of meanings in a lexicon (see Geeraerts [2010] on onomasiology vs. semasiology). Language contact may have other effects on the lexicon: relexification is the result of extensive core borrowing, to the point where a native vocabulary is almost entirely replaced by a foreign one. And this brings us to the notion of degrees of intensity of contact, and Thomason and Kaufman's (1988) much-quoted (but problematic) five-point scale of borrowing: from lexical ('casual' and cultural, then 'regular' and core), to slight structural, to moderate structural, and finally, to heavy structural borrowing.

So lexical borrowing may not be the most dramatic effect of language contact, but it is certainly the most frequent one, and the most productive, resulting in lexical innovation and vocabulary expansion, or what Winters (2015) calls 'lexical layers': sub-lexica in a language whose distinction can add a great deal to the social, historical, and typological study of language. She goes on to identify various types of such layers: armed and political conquest; cultural conquest; religion; adstrata (coexistence of languages due to immigration); and fashion.

3. Courses

Loanwords are often integrated morphologically (as well as phonologically) into the existing structures of the receiving language, often independent of the structures of the source languages. In other words, nativisation processes, independent of the origin of the loans in question (geographical or historical), may be expectable, and even predictable, to a degree. Phonological processes, such as vowel harmonic assimilation, or morphological ones, like nominal suffixation, would tend to recur in the nativisation of words from different languages. A number of studies have focused on the notion of predictability and the general degree of nativisation (e.g. Varella 2006).

In general, the degree of adaptation depends on the degree of similarity between the structures of the systems involved. That is, when the structures are sufficiently similar, the introduction of foreign words causes little or no disturbance to the borrowing language's phonology and morphology, so loans may be adopted without any morphophonological modification. On the other hand, adaptation will be more extreme when the languages involved exhibit major typological differences, or when they possess very divergent phonotactic restrictions and syllable structure constraints. To illustrate, Old

Norse *steik* gave English 'steak', which in turn gave Japanese *suteki*; and French *crème* gave English 'cream' which in turn gave Japanese *curimu* − examples showing more complex adaptation of English loans into Japanese syllable structure.

While the typological distance between languages in contact is a major determinant of the degree of structural adaptation of borrowed vocabulary (see also Chapters 2, 3, 4 and 13), it is less obvious whether there is any systematic correlation between interlinguistic differences and the predictability of the particular processes involved in morphological integration. The latter would be determined by historical processes inherent in the structure of the recipient language. Greek, for example, drew on its whole affixation system to integrate loans from various languages, so that each nominal and verbal group would contain loans from any language. There are exceptions though: (a) a phonetically determined class assignment, e.g. Latin masculine nouns in *-us* forming Greek masculine nouns in *-os* (*campus* > *kámbos* 'plain'), or Latin (and Italian) feminine nouns in *-a* forming Greek nouns in *-α* (*barca* > *várka* 'boat'), or even Turkish nouns in *-i/-ı* forming masculine nouns in *-is* if animate (*boyacı* > *bojadzís* 'painter') or neuter in *-i* if inanimate (*halı* > *xalí* 'carpet'); (b) a semantically determined gender assignment, e.g. Turkish *hasap* > Greek *xasápis* (masc.) 'butcher', Turkish *harup* > Greek *xarúpi* (neut.) 'carob'. In general, what is predictable about morphological adaptation is the range of its possible manifestations − not surprising if we consider that the ultimate, if subconscious, 'goal' of nativisation should be uniformity with the structures of the recipient language (with the exception of often deliberate attempts to reproduce the foreign attributes of certain forms, often leading to a phenomenon known as 'hyperforeignism'; see Janda, Joseph, and Jacobs 1994).

In sum, the predictability of morphological adaptation depends on factors other than the morphological structure of the source language itself: the fact that the Romance languages are marked for both gender and number, whereas English is marked for number only, or the fact that Turkish, unlike other source languages, is an agglutinating one, seemed to have played no role in how the loanwords were treated in Greek. Different factors may play a role in particular processes, like the assignment of gender (e.g. meaning) or class (e.g. phonology; semantic equivalents in the recipient language). Such findings are understandable if we consider that in some cases loans are introduced into a language without the borrower knowing about the structure of the lending language. This is also a reason why a language may borrow words in different forms from the same source, e.g. infinitives treated as stems or inflected forms of verbs undergoing suffix substitution; examples of such processes are furnished by most languages that provided Greek with new lexical material. Generally, how a word sounds or what it means are the important issues in grammatical nativisation. The source grammar matters less: the borrowed inflections of Greek or Latin words in English like 'phenomena' or 'fungi' could invite speculation based on external factors, but counter-examples like 'syllabuses' and 'referendums' make that less feasible.

Another issue relevant to the degree of adaptation concerns the degree of acceptance of loans in the wider borrowing speech community. In general, there seems to be little correlation between the sociolinguistic integration and the structural nativisation of the loanwords. But while the frequent occurrence of some loans in the speech of individuals or groups may not be relevant to their grammatical nativisation, the frequency of borrowing of particular classes of words from particular languages may pertain indeed to their grammatical rendering. Such was the case when particular classes of Provençal and

Venetian or Italian verbs were regarded as prototypical and influenced the nativisation of French and English verbs respectively, providing Greek with productive morphological elements. To illustrate, Venetian or Italian verbs ending in *-are* were adapted in Greek with a phonetically similar suffix *-áro*, e.g. *pontare* > *pondáro* 'to bet'; *voltare* > *voltáro* 'to go for a stroll/ ride'; *provare* > *prováro* 'to try on (clothes)'. There are over a hundred of such verbs, and the suffix *-áro* provided the pattern for the formation of other loan verbs, as in more recent loans from English: 'relax' > *rilaksáro*; 'scan' > *skanáro*; 'surf' > *serfáro*; 'sniff' > *snifáro*; 'speed' > *spidáro*; 'spin' > *spináro*; 'check' > *tsekáro*; and a few more.

It has been argued (Heath 1984; McMahon 1994) that adaptation may follow routines, that is, speakers may adhere to particular methods of borrowing and nativisation, e.g. the addition of a particular suffix to create a verb, which were essentially analogical, based on earlier, similar borrowings from the same or different source, as shown in the previous paragraph. McMahon (1994: 207–208) cites several examples of strategies observed in the grammatical nativisation of loans. Verbalization of a nominal or adjectival form by the addition of a light verb like 'do' is quite common (also in Hock 1991: 386). As for nouns, fitting into the weak or unmarked class in the recipient language, reinterpretation of a plural form as singular, and one gender class becoming the unmarked option for loans, are among attested strategies.

4. Consequences

At any one time, a culture may be absorbing objects and ideas, and the language may be incorporating the names along with these new imports. While this seems to be the norm, there are occasions where the language would rely on its own stock to name new notions, instead of or along with borrowing. Mainly in written language, you could occasionally come across loans and coinages from native material used side-by-side in a way complementary to each other, i.e. one term is cited in order to explain the other. In Greek, the terms 'browser' or 'web browser' are often appositioned to *filometritís istoselíδon*, or *ploiγós web,* or *próγrama periíjisis web*, or *periijitís istú* − terms rarely used in oral speech.

Drawing on native material to create new vocabulary often requires metalinguistic awareness which would be absent from cases of casual cultural encounters and subsequent borrowing. In Greek, the practice of creating compounds was favored almost exclusively by the advocates of *katharevousa* − the attempt to rid Greek of foreign and colloquial influences, particularly in the nineteenth century. It was then that enthusiastic scholars rummaged through classical dictionaries and manufactured a number of terms: so, the common word *ombréla* ('umbrella') became *aleksivróxion* and *aleksílion* (literally 'against the rain' and 'against the sun'), while *patáta* ('potato'), turned into *yeómilon* (literally 'earth-apple', a configuration found across European languages). *Katharevousa* wasn't to last, and neither were these terms.

Similar tendencies are found in places where conservative attitudes and nation-centric ideologies may override what would be the natural process elsewhere. This is how French *ordinateur* and Icelandic *tölva* became established, rather than the near-global 'computer'. But even where direct linguistic contact is absent, the urgent need of a

speech community to name a previously unknown phenomenon will lead to a foreign source. “The designative inadequacy of a vocabulary in naming new things” is, for Weinreich ([1953] 1968: 57), the major reason for lexical borrowing. “Using ready-made designations is more economical than describing things afresh. Few users of language are poets”, he writes. The example of English ‘mass-importing’ Latin and Greek terms during the Renaissance period would be illustrative, where the ‘new’ notions and ideas that emerged from the study of classical literature were henceforth referred to with their foreign names.

Whereas the phenomenon described above involves ‘necessary’ lexical borrowing, as an immediate consequence of cultural borrowing, the second effect of linguistic contact on the vocabulary of a language concerns stylistic additions, i.e. lexical innovations with pre-existent semantic content which complement the native vocabulary and produce doublets in the lexicon. The close contact affecting a language leaves an imprint covering all aspects of everyday life. The composite character of the English lexicon was the consequence, on the one hand, of the existence of bilingual speakers during certain historical periods, including any combination of English, Celtic, Latin, French and Scandinavian. On the other hand, it was the consequence of the communication and interchange of monolingual speakers in what was geographically and politically a single community. “Lexical borrowing requires very restricted bilingualism” (McMahon 1994: 204). This latter type of borrowing is also what happened following the Norman Conquest, where the majority of the native population remained English monolinguals, yet an enormous number of terms connected with administration, warfare, the arts, and the Church, as well as everyday life, were borrowed during the period.

Consequently, whereas ‘necessary’ lexical borrowing tends to involve new terminology of new items, usages, and institutions, loanwords meeting synonymous terms in the borrowing language may end up as stylistic additions. Such doublets in English and other languages are often words of common usage, of high frequency in speech, and which express material and non-material senses, including mainly concrete and abstract nouns, and also some verbs. The latter are, as Whitney (1881) first argued, among the less likely of the vocabulary items to be borrowed by a language, nouns and adjectives coming first. This view is still generally taken to be valid (see the definition of ‘borrowability’ in Trask [2000: 44]).

While it is fair to say that certain lexical categories are more borrowable than others, it would be difficult to make comparisons within a particular category, say, nouns. ‘Basic vocabulary’, a list drawn by Morris Swadesh (original 1950; final revision 1971), was assumed until recently to be the least likely to be affected by lexical innovation, internal or external. However, all proposed categories (see Trask 1996: 22), including body parts, weather, universal experiences, natural phenomena, basic color terms, and kinship terms, have been borrowed into a number of languages (see discussions of Scandinavian loans in English in Lehmann [1992: 269] and Miller [2012: 106–119]). This was Thomason’s (2001) view, supported by Haspelmath and Tadmor (2009), Tadmor (2009), and Tadmor, Haspelmath, and Taylor (2010), who brought extensive cross-linguistic empirical evidence to the debate.

‘Common’ or ‘basic’, and in any case non-cultural, loans are displayed by most languages. A relatively high percentage of loanwords in Greek, for example, cannot be attributed to cultural borrowing. These include abstract nouns (denoting non-intellectual concepts, ideas, feelings, and states-of-being), as well as concrete ones, both frequently

used in speech, and which, at the time of borrowing, could have been the result of individual bilingualism, or even code-mixing, in conversation. Another reason concerns the fact that, at least for literary purposes or for enrichment of style, commonly used words may call for synonyms; that is, authors and poets could have tried to variegate their vocabulary by appealing to some available foreign tongue; or simply they could have been addressing, if not a bilingual population, a community composed of both Greeks and foreigners.

"The unifying factor underlying all borrowing is that of projected gain," Winter (1973: 138) writes, so that the borrowing speech community stands to benefit in some way from the transfer of linguistic material. Beyond the necessity involved in cultural borrowing, such gain may be social, motivated by folk perceptions of prestige and linguistically manifested social aspiration, or it may be more centrally linguistic, since bilingual speakers may find a replacement in their second language for words that seem to be losing their expressive force, have become obscure, and are infrequently used. Moreover, it might be that certain forms lend themselves to borrowing. Johanson (2002) talks about 'attractivity' of forms, i.e. the predisposition of certain forms to transfer. "Linguistic elements may be attractive for speakers because of particular patterning, relatively simple, regular and transparent structures or other properties that make them easy to use and understand" (Johanson 2002: 309). Whether for greater referential potential, or for ease of articulation or comprehension, bilingual speakers are agents of lexical innovation. Especially 'active', widespread bilingualism can lead to massive borrowing, or even to language shift in substrate-type settings. Maltese, for instance, has borrowed so much from Italian and English that its Semitic roots are not immediately manifest; or the case of Albanian, which has lost as much as 90 % of its native vocabulary, after borrowings from Slavic, Hungarian, Greek, Italian and Turkish.

Substantial vocabulary expansion often entails lexical duplicates, which in turn entail stylistic contrasts; that is, either the loanword or the native lexical item may be marked for style, usage in particular contexts, or by some parts of the population. An illustrative example is found in the history of English, and the famous 'inkhorn controversy' of the sixteenth and seventeenth centuries. The new technology of printing, which facilitated wide dissemination of classical literature, resulted in unprecedentedly widespread borrowing. Opponents of the new loans called them 'inkhorn terms' to condemn their literary character and to indicate their obscurity for the majority of English speakers. While words like *fatigate*, *illecebrous*, and *splendidious* did not survive into modern English, many others did: 'logic', 'conclusion', 'negation', and 'affirmation' outlived the then-proposed substitutes *witcraft*, *endsay*, *naysay*, and *yeasay*, respectively (Blank 2006).

Weinreich (1968) summarizes the transfer or reproduction of loanwords with no new content (i.e. non-cultural borrowing) as affecting the existing vocabulary in one of three ways: (a) confusion between the old and the new word; (b) disappearance of the old word; and (c) survival of both the old and the new word, with a specialization in content.

However, other questions emerge: To what extent have the borrowed words managed to become fully nativised and indistinguishable from the native stock? Was 'necessary' cultural borrowing a permanent addition to and an enrichment of the native vocabulary, or merely an import carrying an 'expiry date'? In the case of 'stylistic innovations' or doublets, what was the outcome of a plausible competition between the two? Were the foreign additions to expunge, with their adoption, the native material, or quite the reverse? Would their fate depend on their perceived 'foreignness' or 'familiarity', or on

their degree of integration? In other words, should there be any correlation between the structural and the sociolinguistic integration of loans? Johanson (2002: 300) says: "These processes of integration or nativisation are, in principle, independent of each other and do not necessarily run parallel. Nevertheless, further research may reveal deeper systematic correlations between structural and extra-structural developmental stages."

My observations may indeed suggest correlation between the two phenomena, at least when they are minimally or not at all manifested (since structurally treated vocabulary has also proved to be of restricted and ephemeral use). Such a correlation is easily observable in the case of Greek, due to its rather complex affixation system: words lacking affixes are prone to remain somehow 'outside' Greek, that is, to be socially marked and perceived as foreign.

Looking at Greek historically, we see that cultural loans come and go. As one foreign ruler was replaced by another, and political and social institutions changed, what disappeared was not only isolated terminology, but whole semantic fields, which include all the administrative, legal, and ecclesiastical terms borrowed from Romance languages. Where the Greeks appear to have learned a lot from their conquerors is in the domestic sphere. Terms of architecture, names of household objects, or referring to food, cooking, and clothing have survived, independent of source and date of introduction. In addition, 'no-need' loans also proved long-lasting, widely distributed in the written sources and frequently used in speech. The adoption or disregard of such loanwords might have been dependent on personal preferences; or, generally, on factors far less predictable than in the case of cultural borrowing, whereby the survival of terms was largely dependent on the denotational role they fulfilled.

'No-need' loans may be subject to semantic shift. Some loans show semantic restriction or narrowing. The meaning of the borrowed word becomes less inclusive than the meaning of the original word in the source language. Semantic narrowing is a common process in language history, which is inevitable as the general tendency in lexical innovation is to introduce a new word that would define a concept more precisely when the cultural need arises. That is, cultural or technological advances may necessitate the introduction of hyponyms, i.e. of subordinate terms, for the specialized developments of extant concepts. In language contact, or when a language's own morphological resources (such as derivation or coining) are not used, the specialized term may be a loanword. English, for example, has relatively recently borrowed French *gâteau* ('cake') to denote a rich cake containing layers of cream or fruit; or Greek *trauma* ('injury') to denote a deeply distressing or disturbing experience, or an emotional shock.

The meaning of a hyponymic term may indeed depend on the donor culture, that is to say, the adaptation of the word's original specification may be culturally or contextually triggered: English speakers use the term *perestroika* referring only to 'Russian political restructuring', though it is a general term for 'restructuring' in the donor language; also general is *chansonnier* in French, which is used in English to denote the 'singer of French ballads'.

Less expected semantic shifts may also be observed. The 'new' meaning may be largely, or only, dependent on the borrowing speaker's knowledge or receptivity − a phenomenon similar to the transition of foreign phonology, where unfamiliar sounds are likely to be modified accordingly. Or it may be exclusively bound to the collocational or circumstantial context in which the word is used. Therefore, even in cases where speakers of one language using a particular word still have a referent at their disposal,

the context given to the foreign speakers may be inadequate during cross-linguistic exchange. Under such circumstances, conclusions of the kind 'what we hear is what we see' can be erroneous. Unsuccessful connections between the words and the referents can often occur.

Patterns that would provide some insight as to how the speakers of the borrowing speech community go about assigning meaning to the foreign words they encounter can be identified, and include: (a) displacements of meaning within the same semantic field, e.g. Latin *candela* 'candle' > Greek *kandíli* 'lamb'; (b) transfer to a different semantic field, due to similarity of percepts or other semantic associations, e.g. Latin *pluma* 'feather' > Greek *plumí* 'jewel' ; or (c) adaptation of meaning derived from an overt manifestation of the abstract sense of the original word, e.g. Venetian *spazzaura* 'trash' > Greek *patsavúra* 'rag'. Some adaptations are simply metaphorical renditions of the literal senses of the source words: 'ecstasy' comes from Greek *ékstasis* 'standing outside oneself'; 'sarcasm' from Greek *sarkázo* 'to bite one's lips'. It's possible that the etymological roots may conjure up visual images which could suggest how the English words mean what they do.

In general, apparently curious shifts of meaning can be logically grounded. In addition, they can be verified at a psycholinguistic level, in particular by relevant findings connecting the notions of concreteness, imageability, and use. It has been argued, for example, that concrete words are more imageable than abstract words, that they are more easily recalled in memory, and that they occur more frequently in word-access (Reeves, Hirsh-Pasek, and Golinkoff 1998: 167−168).

5. Conclusions

Psycholinguistics may hold answers to issues we encounter in the study of lexicogenesis and language contact. The link is an obvious one: lexical borrowings often originate in bilingual speech as nonce-formations, fleeting code-switches with the potential to be repeated, copied, and spread. Much has been said about the necessity or otherwise of a distinction between lexical borrowing and code-switching (Appel and Muysken 1987; Myers-Scotton 1994). More recent studies combine synchronic and diachronic approaches: Winford (2002) considers not only code-switching, but also second language acquisition; Matras (2009) looks at bilingual acquisition, the bilingual individual, and societal multilingualism.

Cognitive linguistics may contribute to the study of language contact and word formation. Grondelaers, Speelman, and Geeraerts (2007) explain how central ideas in cognitive semantics, such as prototype theory and categorization, can explain not only semasiological but also onomasiological change. *Computer* in German represents both a conceptual and lexical innovation. How did this conceptualization take place? Moreover, what does the coexistence or competition between the native word *Rechner* and the borrowed *Computer* suggest about the referential (denotational) and non-referential (connotational) meaning of the words? Studies of cognition may contribute to issues such as borrowing versus morphological productivity, and to questions on the universality of conceptualizations.

Finally, Hanks' (2013) Theory of Norms and Exploitations attempts to explain creative language use − 'exploitation' of the 'normal', central, and typical semantic and

collocational properties of words. Borrowing is of course one of the main creative processes in language, and Hanks illustrates with some Shakespearean formations what can be done by 'exploiting' foreign lexis, drawing on corpora for deciding the 'norms' and following the typology of loans discussed in the first part of this chapter.

6. References

Appel, René & Pieter Muysken
1987 *Language contact and bilingualism*. London: Arnold.
Blank, Paula
2006 The Babel of Renaissance English. In Lynda Mugglestone (ed.), *The Oxford history of English*, 212–239. Oxford: Oxford University Press.
Durkin, Philip
2006 Lexical borrowing in present-day English. *Oxford University Working Papers in Linguistics, Philology and Phonetics* 11. 26–42.
Durkin, Philip
2009 *The Oxford guide to etymology*. Oxford: Oxford University Press.
Fischer, Andreas
2003 Lexical borrowing and the history of English: A typology of typologies. In Dieter Kastovsky & Arthur Mettinger (eds.), *Language contact in the history of English*, 97–115. Frankfurt: Peter Lang.
Geeraerts, Dirk
2010 *Theories of lexical semantics*. Oxford: Oxford University Press.
Geeraerts, Dirk
2015 How words and vocabularies change. In John R. Taylor (ed.), *The Oxford handbook of the word*, 416–430. Oxford: Oxford University Press.
Grant, Anthony P.
2015 Lexical borrowing. In John R. Taylor (ed.), *The Oxford handbook of the word*, 431–444. Oxford: Oxford University Press.
Grondelaers, Stefan, Dirk Speelman & Dirk Geeraerts
2007 Lexical variation and change. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 988–1011. Oxford: Oxford University Press.
Hancock, Ian F.
1971 A provisional comparison of the English-derived Atlantic creoles. In Dell Hymes (ed.), *Pidginization and creolization of language*, 287–291. Cambridge: Cambridge University Press.
Hanks, Peter
2013 *Lexical analysis: Norms and exploitations*. Cambridge, MA: MIT Press.
Haspelmath, Martin
2009 Lexical borrowing: Concepts and issues. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 35–54. Berlin: Mouton de Gruyter.
Haspelmath, Martin & Uri Tadmor
2009 The loanword typology project and the world loanword database. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 1–34. Berlin: Mouton de Gruyter.
Haugen, Einar
1950 The analysis of linguistic borrowing. *Language* 26(2). 210–231.
Heath, Jeffrey G.
1984 Language contact and language change. *Annual Review of Anthropology* 13. 367–384.

Hock, Hans H.
1991 *Principles of historical linguistics.* Berlin: Mouton de Gruyter.
Janda, Richard D., Brian D. Joseph & Neil G. Jacobs
1994 Systematic hyperforeignisms as maximally external evidence for linguistic rules. In Susan D. Lima, Roberta Corrigan & Gregory K. Iverson (eds.), *The reality of linguistic rules*, 67–92. Amsterdam: Benjamins.
Johanson, Lars
2002 Contact-induced change in a code-copying framework. In Mari C. Jones & Edith Esch (eds.), *Language change: The interplay of internal, external and intra-linguistic factors*, 285–313. Berlin: Mouton de Gruyter.
Lehmann, Winfred P.
1992 *Historical linguistics*. London: Routledge.
Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.
McMahon, April
1994 *Understanding language change.* Cambridge: Cambridge University Press.
Miller, D. Garry
2012 *External influences on English: From its beginnings to the Renaissance.* Oxford: Oxford University Press.
Miller, D. Garry
2014 *English lexicogenesis*. Oxford: Oxford University Press.
Milroy, James
1989 The concept of prestige in sociolinguistic argumentation. *York Papers in Linguistics* 13. 215–226.
Myers-Scotton, Carol
1994 *Lexical borrowing: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
Reeves, Lauretta, Kathy Hirsh–Pasek & Roberta Golinkoff
1998 Words and meaning: From primitives to complex organization. In Jean Berko Gleason & Nan Bernstein Ratner (eds.), *Psycholinguistics*, 157–226. Belmond CA: Wadsworth.
Swadesh, Morris
1950 Salish Internal Relationships. *International Journal of American Linguistics* 16(4). 157–167.
Swadesh, Morris
1971 *The origin and diversification of language* (ed. by Joel Sherzer). Chicago: Aldine Atherton.
Tadmor, Uri
2009 Loanwords in the world's languages: Findings and results. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 55–75. Berlin: Mouton de Gruyter.
Tadmor, Uri, Martin Haspelmath & Bradley Taylor
2010 Borrowability and the notion of basic vocabulary. *Diachronica* 27(2). 226–246.
Thomason, Sarah G.
2001 *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
Thomason, Sarah G. & Terence Kaufman
1988 *Language contact, creolization and genetic linguistics*. Berkeley, CA: University of California Press.
Trask, Robert Lawrence
1996 *Historical linguistics*. London: Arnold.
Trask, Robert Lawrence
2000 *The dictionary of historical and comparative linguistics*. Edinburgh: Edinburgh University Press.

Varella, Stavroula
2006 *Language contact and the lexicon in the history of Cypriot Greek*. Bern: Peter Lang.
Weinreich, Uriel
1968 [1953] *Languages in contact: Findings and problems*. The Hague: Mouton.
Whitney, William D.
1881 On mixture in language. *Transactions of the American Philological Association* 12. 1–26.
Winford, Donald
2002 *An introduction to contact linguistics.* Oxford: Blackwell.
Winter, Werner
1973 Areal linguistics: Some general considerations. In Thomas A. Sebeok (ed.), *Current trends in linguistics II: Diachronic, areal and typological linguistics*, 135–147. The Hague: Mouton.
Winters, Margaret E.
2015 Lexical layers. In John R. Taylor (ed.), *The Oxford handbook of the word*, 445–461. Oxford: Oxford University Press.

Stavroula Varella, Chichester (UK)

6. Morphology

1. Introduction
2. Borrowing
3. Insertional code-switching
4. Mixed languages
5. Pidgin and creole languages
6. Conclusion
7. References

1. Introduction

The transfer of morphology is of interest to the field of language contact because it occurs less frequently than lexical transfer. Morphology is borrowed less often than vocabulary (Section 2), and is generally only derived from the dominant of the two interacting languages in insertional code-switching (Section 3). In contrast, morphology from both source languages is maintained in some mixed languages (Section 4). Finally it is argued that pidgin and creole languages contain relatively little morphology compared with their lexifier languages, although this claim is controversial (Section 5).

These generalizations about morphology apply to varying degrees to different types of morphology. Derivational morphology is generally more likely to undergo transfer or be maintained than inflectional morphology, and inherent inflection is generally more resilient than contextual inflection. This borrowability hierarchy is based on the degree to which a morpheme relates to other parts of the clause outside its maximal projection, e.g. contextual inflection, such as agreement markers and case morphology, is exceptionally fragile in contact situations due to its syntactically dependent nature.

https://doi.org/10.1515/9783110435351-006

Given the sensitivity of morphology to language contact, it can be used as a litmus test to gauge the relative strengths of interacting languages. For example, in borrowing and code-switching, one language is more dominant, as defined by the presence of inflectional morphology. On the other hand, the maintenance of inflectional morphology from both languages in some mixed languages suggests a relatively equal weighting given to both languages, with neither language being definitively stronger.

2. Borrowing

The status of morphology in borrowing has received much attention due to the apparent difficulty in transferring morphology, particularly inflections (Aikhenvald and Dixon 2007; Matras and Sakel 2007; see also Chapter 13). This difficulty is shown most clearly through borrowability hierarchies which are based on cross-linguistic surveys of languages in contact. Borrowability hierarchies are scales of lexical and grammatical categories which are more or less likely to be transferred. A number of explanations have been proffered to explain these cross-linguistic tendencies. Structural constraints are said to affect the ability of morphemes to transfer from one language to another. These constraints are based on the typology of the languages in contact (Heath 1978; Weinreich [1953] 1974), or on the degree of syntactic involvement a morpheme has in the clause (Gardani 2008). On the other hand, Thomason and Kaufman (1988) suggest that any morpheme may be borrowed, given the right level of social disruption, with the intensity of the contact situation determining the degree of borrowing. Note that these borrowability hierarchies relate to the transfer of the phonological form of a morpheme, called MAT(ter) borrowing by Sakel (2007) (Section 2.1). Less attention has been given to the transfer of structural patterns called PAT(tern) borrowing (Section 2.2).

2.1. MAT(ter) borrowing

The study of borrowing patterns and constraints began with Whitney (1881), who created a hierarchy of borrowing according to grammatical categories. Nouns were considered the most susceptible to borrowing, followed by other parts of speech, suffixes, and inflections. In this scale, Whitney (1881) did not preclude the borrowing of inflectional morphology, however he did suggest that it was extremely unlikely. Similar views were expressed later. In particular, Haugen (1950: 225) conducted a study of borrowing in American Norwegian and American Swedish, and concluded that "the more structural a feature is, the less likely it is to be borrowed". Singh's (1982) study of English borrowings into Hindi also produced a similar hierarchy, as did Muysken's (1981) study of Spanish borrowings into Quechua.

Based on these descriptive hierarchies, Moravcsik (1978: 110–12) posits an implicational hierarchy of borrowing. She suggests that non-lexical items will not be borrowed unless some lexical items have already been borrowed (lexical > functional), borrowed lexical items such as verbs will only be observed in a language if borrowed nouns are already present (nouns > other lexical items), and that "no inflectional affixes can belong

to the set of properties borrowed from a language unless at least one derivational affix also belongs to the set" (derivational > inflectional).

Of course, all of these hierarchies are probabilistic and do not categorically preclude the transfer of morphology. A number of other case studies of derivational borrowings can be found in Gardani, Arkadiev, and Amiridze (2015), and have also been collated in Seifart's AfBo database (http://afbo.info). Inherent inflections have also been observed. For example, it is not uncommon to find cases of borrowed plural markers which have been extended to non-native words. Gardani (2008) gives a number of examples of borrowings of this type, such as Turkish plural marking into Albanian, Greek into Aromunian, and Arabic into New Persian. Other examples of inflectional borrowings can be found in Thomason (2015) and Clements and Luís (2015). A number of instances of contextual borrowings are also present in the literature. Meakins (2015) observes that Australia has a high number of case borrowings (e.g. Heath 1978), which she attributes to the non-argument status of nouns in non-configurational languages (Meakins 2011). These nouns operate at the morpho-syntactic periphery of clauses, acting as 'Trojan horses' to smuggle inflections into a recipient language.

Paradigm borrowing is the only inflectional category which appears to be almost completely resistant to borrowing; however some counterclaims exist. For example, Seifart (2012) has described the wholesale transfer of the full paradigm of noun class suffixes from Bora (Isolate, Brazil) into Risígaro (Arawakan, Brazil). Similarly, Sebjan-Küöl Even (Siberia) has borrowed the Sakha assertive-presumptive mood paradigm, plus associated subject agreement suffixes, without their native verb stems (Pakendorf 2009).

Although inflectional borrowings have been observed, their relative rarity requires some explanation. Weinreich (1974: 32−35) provides mechanisms whereby inflectional morphemes may be transferred. He suggests that if the structures of the source and recipient languages are congruent, then transfer is strongly facilitated. Overt morphemes also tend to replace zero morphemes (Gardani 2008: 88), and relatively unbound morphemes are more likely to replace a bound counterpart in another language. Morphemes with complex functions, e.g. inflections, are also less likely to be borrowed, and the more salient the morpheme's boundary is, e.g. agglutinating morphology, the easier it is to borrow; see also Heath's (1978: 105−07) suggestion of alternative factors which affect the borrowability of inflectional morphology, and Winford's (2003: 92) reconciliation of both positions.

Two principles seem to be at work in Weinreich's (and Heath's) accounts: (i) the compatibility or congruence of both languages in the borrowing relationship, and (ii) the morpheme's degree of structural integration in the clause. Firstly, structural similarity between languages has been shown to facilitate transfer (see Law [2013] for a recent example; and Sebba [1998] for a similar claim in the code-switching literature). Secondly, Gardani (2008) uses Booij's (1996) distinction between contextual and inherent inflection to highlight the role of structural integration in borrowing. Contextual inflection is determined by the syntactic contexts in which it occurs. Structural case and subject-verb agreement, which mark the relationship between a head and a dependent, are good examples. On the other hand, inherent inflection is determined by the information a speaker wishes to convey, and is exemplified by plural affixes, local case markers (such as the locative and allative), and TAM inflection. Gardani (2008) suggests that inherent inflection (and derivation) is more likely to be borrowed than contextual morphology because it is not dependent on other parts of the grammar. The use of Booij's (1996)

distinction provides a probabilistic explanation for why the borrowing of inherent morphology, such as plural markers, occurs more often than case marking.

Thomason and Kaufman (1988) provide an alternative explanatory model of borrowing based on social factors, believing that all structures can be borrowed regardless of the structural category and the typological distance between the affected languages. They suggest that social factors play the determinate role in the linguistic outcome of language contact, and that, given the right level of social disruption and bilingualism (Thomason and Kaufman 1988: 33, 47), morphological borrowing is not unusual. They propose a borrowing scale not unlike previous scales; their scale differs in that it is based on the degree of contact rather than structural features. It nonetheless correlates very neatly with previous observations about the degree of borrowing of structural features.

2.2. PAT(tern) borrowing

The type of borrowing discussed so far consists of the transfer of phonological form. PAT(tern) borrowings (see Sakel 2007), i.e. changes to distributional patterns in a clause, also affect morphology; for example, it is not uncommon to find shifts in case alignment as a result of convergence with a case system from a different language. Typically, the case system of one language undergoes realignment under the influence of another language, without a corresponding change to the native case forms. Note that MAT and PAT borrowing are also related; for example, the premise of King (2000) is that lexical borrowing can trigger internal changes in the grammar of a recipient language.

Changes to peripheral case markers are not uncommon. For example, Basque has fourteen case suffixes; however, the comitative suffix *-ekin* is increasingly used to express functions previously expressed by the instrumental case. This change has occurred under the influence of the Spanish preposition *con*, which expresses both functions (Heine and Kuteva 2005: 149–50). A similar change has occurred in Imbabura Quichua, where a distinction between comitative *-wan* and instrumental case *-ntin* has been lost under the influence of Spanish (Gómez Rendón 2007: 486–87). The Russian case system has influenced Evenki (Tungusic), where the agent of the passive is now marked with the instrumental rather than the dative (Grenoble 2000: 109–10). Kupwar Kannada (Dravidian) has extended its dative postposition from marking not only human direct objects, but also to human indirect objects, under the influence of Urdu and Marathi (Indo-Aryan) (Gumperz and Wilson 1971: 158).

Changes to core case categories in situations of language contact are less common than shifts in the alignment of peripheral case categories. Such changes result in the complete morpho-syntactic realignment of grammatical relations. For example, Ardeşen Laz (Kartvelian, Turkey), which had a split ergative system, lost all of its core case markers under the influence of Turkish and developed a neutral system of case alignment (Kutscher 2001). Reverse shifts from accusative to ergative systems have also been observed. For example, in Anatolia, Hittite (Indo-European) developed a split ergative system as a result of contact with Hurrian (Hurro-Urartian), which is an ergative language (Garrett 1990).

3. Insertional code-switching

Insertional code-switching is another area of language contact where different types of morphemes exhibit different patterns of distribution (see also Chapter 14). This type of code-switching involves introducing elements from one language into the morpho-syntactic frame of another, sometimes called the 'matrix language' (Muysken 2000; Myers-Scotton 2002). Two main approaches to this type of code-switching exist – social motivations accounts (Auer 1998; Fishman 1972; Gumperz 1982) and structural constraints theories. Of interest here are structural accounts which examine the shape of code-switching, and the restrictions on mixing grammars.

Inflectional morphology is accorded a special place in studies of code-switching. This type of morphology is often used to identify the matrix language (e.g. Myers-Scotton 2002; Treffers-Daller 1994), and other proposals for structural constraints on code-switching are then based on the identification of the matrix language. Myers-Scotton's (2002) Matrix Language Frame (MLF) model of code-switching deals specifically with the behavior of different types of morphemes in code-switching, and in particular, with inflectional morphology.

The MLF model is based on a unique categorization of morphemes called the 4-M Model. Morphemes are first divided into *content* and *system* morphemes, with system morphemes further divided into *early* and *late* system morphemes. Late system morphemes are of two types: *bridge* and *outsider* morphemes.

Content morphemes, e.g. nouns and verbs, participate in the thematic grid of the utterance. They assign or receive thematic roles, where system morphemes do not. *System* morphemes are more functional. *Early* system morphemes do not assign or receive thematic roles; however they pattern with the content, adding extra meaning to the head of a phrase. These morphemes, e.g. determiners and plural markers, also depend on the head (a content morpheme) of their maximal projection for their syntactic role. *Late* system morphemes do not convey conceptual information, but rather grammatical information. They are structurally assigned outside of their maximal projections to indicate relations between elements in the Complementizer Phrase rather than lower level phrases. There are two different types of late system morphemes: *bridge* system morphemes and *outsider* system morphemes. The difference between these two morphemes lies in where they receive their assignment. Bridge system morphemes depend on information from *within* their maximal projection, whereas outsider system morphemes rely on a source *outside* of their immediate maximal projections. Examples from English of *bridge*

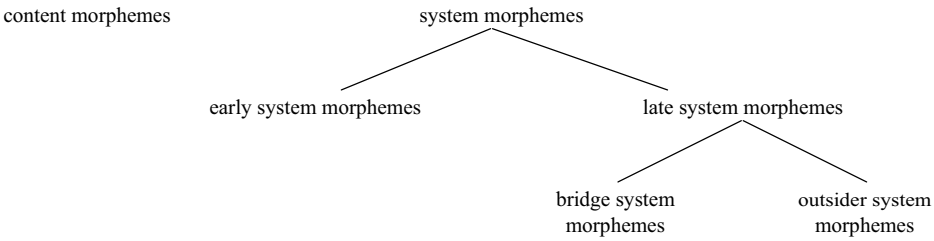


Fig. 6.1: Myers-Scotton's (2002) 4-M Model

system morphemes are the expletive *it*, and *of* in possessive constructions, such as *the foot of the hill*. Outsider system morphemes include contextual inflection, for example, subject-verb agreement markers and case morphology.

Within this morphological framework, Myers-Scotton (1993: 83) predicts that insertions into a matrix language are largely noun and verb switches, and late system morphemes are blocked from switching because they rely on the larger syntactic context. This prediction is called the System Morpheme Principle and is largely upheld by data. Parallels can be made with Gardani's (2008) predictions about borrowing inherent vs. contextual inflection (Section 2.1).

Nonetheless, counter examples to Myers-Scotton's (1993) System Morpheme Principle exist. Similar to the borrowing literature, many cases of inserted nouns accompanied by a plural marker can be found. Past and present participle forms are often found because they accompany inserted verbs (Muysken 2000: 173–76). More perplexing are cases of Turkish verb morphology, with Turkish verbs inserted into Romani in Greek Thrace (Adamou 2010), and the use of Gurindji case morphology in Kriol matrix clauses (Meakins 2014). Nonetheless, these examples are rare and if Myers-Scotton's (1993) predictions are treated as probabilistic constraints rather than absolute rules, they account for most code-switching data. Muysken (2000: 25; see also Sebba 1998) notes that the structural congruence of languages in contact often "undoes" constraints, which may account for these aberrant cases.

4. Mixed languages

Most mixed languages develop from code-switching or borrowing (Bakker 2015; Meakins 2013). Although most mixed languages, referred to as L-G mixed languages, combine the grammar of one language with the partial lexicon from another language (and therefore do not involve morphological transfer), other mixed languages maintain large amounts of morphology from both source languages. From the previous sections on the behavior of morphology in situations of borrowing and code-switching, it is obvious that combining morphological systems is quite exceptional (Meakins et al. 2019). Indeed, Matras (2003) suggests that this is a particular feature of mixed languages – the seemingly unconstrained transfer of grammatical elements, such as inflectional morphology, which in the past have been labelled 'loan proof'. Two types of morphological mixes can be found: the first combines morphological forms from both languages (Section 4.1), and the second restructures the morphological forms from one language according to the distributional patterns of another (Section 4.2).

4.1. Structural mixes (MAT transfer)

A small group of mixed languages contains morphology from both languages. These languages include Michif, Mednyj Aleut, Gurindji Kriol, and Light Warlpiri. In these languages, both source languages contribute to the morpho-syntactic frame of the language, creating what Myers-Scotton (2003) calls a 'composite matrix language'.

First, morphology from both French and Cree is present in Michif, which is spoken in Canada. Verbal inflections are derived from Cree, and the nominal system preserves both French plural morphology and adjectival agreement (Bakker 1997). Mednyj Aleut, spoken on Mednyj Island in the Bering Strait, consists of many Aleut nominal inflections, including two case distinctions (absolutive and relative) and various derivational suffixes (agentive, instrumental, locative, detransitive, inchoative markers, etc.), as well as finite verbal inflectional morphology from Russian, including portmanteau morphemes which express tense, number, and person markers (Golovko 1994). Light Warlpiri and Gurindji Kriol, spoken in northern Australia, exhibit a split between the nominal and the verbal systems. The nominal structure, including case marking (ergative, dative, locative, allative, and ablative) is from the heritage language, Warlpiri or Gurindji, with Kriol providing the verbal structure, including TAM auxiliaries and some suffixes (Meakins 2011; O'Shannessy 2013).

4.2. Converted languages (PAT transfer)

Converted languages develop when the ancestral language maintains its lexical and morphological forms (MAT), but undergoes a complete restructuring of its morpho-syntax (PAT) as a result of contact with an introduced language. They form as a result of a process called *metatypy* (Ross 2006). Sri Lanka Malay is a Malay/Indonesian (Austronesian) variety heavily restructured under the influence of Tamil (Dravidian), and more recently of Sinhala (Indo-Aryan), which occurred as a result of sustained social contact with Tamil-speaking Moors and pervasive Malay-Tamil bilingualism among Malay descendants. The result is a language which is unintelligible to Malay speakers, despite its Austronesian lexicon. Structurally, Sri Lanka Malay developed from an isolating language to an agglutinating language under the influence of Tamil. It has also acquired SOV word order, postpositions, and pre-nominal determiners and adjectives due to this contact (see papers and references therein from Nordhoff [2012]). In terms of morphology, Sri Lanka Malay developed a set of case suffixes from Vehicular Malay prepositions under the influence of Tamil post-posed case-markers. Moreover, even though the forms are from Vehicular Malay, they reflect the patterns of marking found in Tamil. For example, the Sri Lanka Malay-derived locative marker *-ka* marks both location and goal, following the pattern of Tamil (Smith, Paauw, and Hussainmiya 2004: 206).

5. Pidgin and creole languages

The status of morphology in pidgin and creole languages has been the subject of ongoing debate, and relates to a similar argument about whether pidgins and creoles constitute a separate typological class of language (see also Chapter 9). Most introductory textbooks contain cursory notes about the paucity of morphology in these languages, often based on the work of Bickerton (1981). But, as DeGraff (2001: 55) and Bakker (2003: 3) note, this type of claim can be traced back as far as the 1920s, and was usually accompanied by descriptions of the speakers as naïve and simple (Hjelmslev 1938; Jespersen 1922). More recently, the discussion about the status of morphology in these languages was

resurrected with McWhorter's (1998) Creole Prototype. McWhorter (1998) suggested that a universal feature of creole languages is their lack of inflectional morphology and non-compositional derivational morphology. He claimed that the absence of these features is not accidental, as they are related to "definable signs of youth in the structure of a new language" (McWorther 2005: 10), i.e. the typological life cycle of morphology begins with free morphemes which may become morphology over time, an observation deriving from grammaticalization theory.

As a consequence of this and earlier proposals, a number of surveys of inflectional morphology in these languages ensued (Plag 2003a, 2003b; Roberts and Bresnan 2008), which then generated counter surveys (Bakker et al. 2011; Parkvall 2008). These surveys showed that morphology is found to be present, but scarce, in these languages, although the inflectional categories are often present, expressed as free morphemes such as auxiliaries (Kihm 2003). The publication of the Atlas of Pidgin and Creole Language Structures (APiCS) provides more scope for such surveys: http://apics-online.info (Michaelis et al. 2013).

In terms of verb morphology, Kitúba (Zaire, Congo, and Angola) is the one creole language which has tense morphology. It has a tense suffix − anterior *-a(k)a* (with zero marking denoting narrative tense) − which is derived from tense prefixes in its lexifier language, Kikóngo (Bantu) (Mufwene 1997: 179). Aspect morphology is slightly more common than tense morphology, noted in a number of pidgin and creole languages. Berbice Dutch Creole has three aspect suffixes from its substrate language Eastern Ijo − perfective, imperfect, iterative (Kouwenberg 1994) − and Australian Kriol contains a progressive suffix *-bat* derived from the English preposition *about* (Munro 2004: 88). Bakker (2003: 19) notes that Korlai, a Portuguese lexifier creole has both tense and aspect suffixes, maintaining a distinction between a plain form, gerund, simple past, and perfect. Bakker (2003: 21) summarizes his survey of pidgin languages in an extensive table of verbal morphology. For example, he finds tense suffixes in Lingua Franca (Pidgin Italian), Fanakalo (Pidgin Zulu) (also Bakker 1995: 33), and Bilkiire (Pidgin Fula), and aspectual prefixes in Pidgin Ojibwe.

Nominal inflectional morphology in the form of number, gender, and word class can also be observed in pidgin and creole languages, although case morphology is extremely rare. Plural markers have been noted in a number of pidgin languages, including Turku (pidgin Arabic found in Chad) and Kenyan Pidgin Swahili (Bakker 2003: 15). Baptista (2003: 317, 323, 324) also observes plural markers in a type of Cape Verdean Creole, Palenquero (Spanish lexifier) and Ghanian Pidgin English. Gender morphology is even more rarely reported. Bakker (2003: 16) notes that gender is preserved to some extent in pidgins with Bantu lexifiers, for instance, Kenyan Pidgin Swahili and Fanakalo, which is not surprising, given that Bantu languages are known for their sophisticated noun class systems. DeGraff (2001: 73) finds gender marking in Haitian (a French lexifier creole); however, as Kihm (2003: 345) notes, only *natural* gender is marked, which is clearly different from French, where every noun is assigned a gender value. Kitúba (Kikóngo lexifier) also seems to have preserved the noun-class system of its lexifier (Mufwene 1997: 177−78). Finally, case morphology is not found in any pidgin or creole language.

On a whole, these surveys did not find inflectional morphology to be lacking, although it was present in smaller quantities than in their lexifier languages. Of course, one of the problems with this observation is that most of the lexifier languages of attested creoles (e.g. English, French, Dutch) contain relatively little inflectional morphology

themselves. However, there are a couple of examples of creoles with highly inflected lexifiers where claims about the transfer of inflectional morphology can be tested properly. For example, Kitúba has an agglutinating language as its lexifier language, Kikóngo (Bantu) (Mufwene 1997). Kikóngo has an elaborate noun class and bound pronoun system, including subject-verb agreement, and a verbal tense/aspect system. Different outcomes can be observed for these systems. Significantly, the noun class system has been preserved, although some nouns have changed class. Subject-verb agreement has been lost completely, and in the case of the bound pronouns and TAM suffixes, the grammatical categories of Kikóngo have been preserved, however they are expressed using free forms in Kitúba (only one verb suffix has been retained) (Mufwene 1997: 175–79). Similarly, Pidgin Ngarluma (Australia) is based on Ngarluma, which is a highly inflected suffixing language, with case-marked nominals (nominative and accusative) and bound verb roots which inflect for TAM categories. Pidgin Ngarluma is markedly simpler. It is devoid of case-marking, and frozen forms of verbs are used (usually present or future/purposive forms). It also lacks a case distinction between pronouns, at least in the second person (Dench 1998: 32–38).

The claims of McWhorter's (1998) Creole Prototype have also been attacked from an ideological angle. In particular, DeGraff (2001: 54–57) and Lefebvre (2004) take issue with McWhorter's general characterization of creoles as simple. DeGraff refers to earlier similar claims made by Seuren (1998) and Whinnom (1971). This criticism must be viewed within a more general debate about the classification of creole languages based on typological characteristics rather than socio-historical features. The former approach forms a part of what DeGraff (2004, 2005) labels 'Creole Exceptionalism'. DeGraff suggests that Creole Exceptionalism is a colonial discourse within academic writings on creole languages which perpetuates the marginalization of these languages and their speakers. It involves the "postulation of exceptional and abnormal characteristics in the diachrony and/or synchrony of creole languages as a class" (DeGraff 2005: 534). DeGraff suggests that the typological category of a creole is in fact a construct of this academic discourse. Given the history of less than favorable writings on creole languages and their speakers, even in quite recent descriptions (Whinnom 1971: 109), this response has merit. A heightened awareness of the discourse used to describe creoles has meant that phrases which are emotionally neutral in typology or even historical linguistics, such as 'the simplification of inflectional morphology', evoke intense feeling. Typological comparisons with so-called 'normal' languages have become highly charged debates as a result, but they continue nonetheless; see Bakker et al. (2011) and Parkvall (2008).

6. Conclusion

Observations on the behavior of morphology in situations of language contact have played a fundamental role in theoretical accounts of contact-induced change. Although not 'loan proof', morphology, and particularly contextual inflection, is highly resistant to transfer both in situations of borrowing and code-switching in comparison with lexical categories. This resistance has provided historical linguistics with new measures of genetic lineage, which promises to push back the 'glass ceiling' date for historical linguis-

tics of 6000 years (e.g. Dunn et al. 2005). In a related vein, computational phylogenetics and population genetics are offering new methods for measuring large-scale borrowing across language families and within populations of speakers (List et al. 2013; Meakins et al. 2019).

7. References

Adamou, Evangelia
2010 Bilingual speech and language ecology in Greek Thrace: Romani and Pomak in contact with Turkish. *Language in Society* 39(2). 147–171.

Aikhenvald, Alexandra Y. & Robert M. W. Dixon
2007 *Grammars in contact: A cross-linguistic typology*. Oxford: Oxford University Press.

Auer, Peter
1998 *Code-switching in conversation: Language, interaction and identity*. London: Routledge.

Bakker, Peter
1995 Pidgins. In Jacques Arends, Pieter Muysken & Norval Smith (eds.), *Pidgins and creoles: An introduction*, 25–40. Amsterdam: John Benjamins.

Bakker, Peter
1997 *A language of our own: The genesis of Michif, the mixed Cree-French language of the Canadian Métis*. New York: Oxford University Press.

Bakker, Peter
2003 Pidgin inflectional morphology and its implications for creole morphology. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 2002*, 3–33. London: Kluwer.

Bakker, Peter
2015 Typology of mixed languages. In Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds.), *The Cambridge handbook of linguistic typology*, 217–253. Cambridge: Cambridge University Press.

Bakker, Peter, Aymeric Daval-Markussen, Mikael Parkvall & Ingo Plag
2011 Creoles are typologically distinct from non-creoles. *Journal of Pidgin and Creole languages* 26(1). 5–42.

Baptista, Marlyse
2003 Inflectional plural marking in pidgins and creoles: A comparative study. In Ingo Plag (ed.), *Phonology and morphology of creole languages*, 315–332. Tübingen: Max Niemeyer Verlag.

Bickerton, Derek
1981 *Roots of language*. Ann Arbor: Karoma.

Booij, Geert
1996 Inherent versus contextual inflection and the split morphology hypothesis. *Yearbook of Morphology 1995*. 1–16.

Clements, Joseph & Ana Luís
2015 Contact intensity and the borrowing of bound morphology in Korlai Indo-Portuguese. In Francesco Gardani, Peter Arkadiev & Nino Amiridze (eds.), *Borrowed morphology*, 219–239. Berlin: Mouton de Gruyter.

DeGraff, Michel
2001 Morphology in creole genesis: Linguistics and ideology. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 53–121. Cambridge, MA: MIT Press.

DeGraff, Michel
2004 Morphology and word order in 'creolization' and beyond. In Giglielmo Cinque & Richard Kayne (eds.), *The Oxford handbook of comparative syntax*, 293–336. New York: Oxford University Press.

DeGraff, Michel
2005 Linguists' most dangerous myth: The fallacy of Creole Exceptionalism. *Language in Society* 34(4). 533–591.
Dench, Alan
1998 Pidgin Ngarluma: An Indigenous contact language in north western Australia. *Journal of Pidgin and Creole languages* 13(1). 1–61.
Dunn, Michael, Angela Terrill, Ger Reesink, Robert Foley & Stephen Levinson
2005 Structural phylogenetics and the reconstruction of ancient language history. *Science* 309(5743). 2072–2075.
Fishman, Joshua A.
1972 Domains and the relationship between micro- and macro- sociolinguistics. In John Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics*, 435–453. New York: Holt, Rinehart and Winston.
Gardani, Francesco
2008 *Borrowing of inflectional morphemes in language contact*. Frankfurt: Peter Lang.
Gardani, Francesco, Peter Arkadiev & Nino Amiridze (eds.)
2015 *Borrowed morphology*. Berlin: De Gruyter Mouton.
Garrett, Andrew
1990 The origin of NP split ergativity. *Language* 66(2). 261–296.
Golovko, Evgeniy
1994 Mednyj Aleut or Copper Island Aleut: An Aleut-Russian mixed language. In Peter Bakker & Maarten Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, 113–121. Amsterdam: Uitgave IFOTT.
Gómez Rendón, Jorge
2007 Grammatical borrowing in Imbabura Quichua. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 491–521. Berlin: Mouton de Gruyter.
Grenoble, Lenore
2000 Morphosyntactic change: The impact of Russian on Evenski. In Dicky Gilbers, John Nerbonne & Jos Schaeken (eds.), *Languages in contact*, 105–120. Amsterdam: Rodopi.
Gumperz, John
1982 *Discourse strategies*. Cambridge: Cambridge University Press.
Gumperz, John & Robert Wilson
1971 Convergence and creolization. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 151–168. Cambridge: Cambridge University Press.
Haugen, Einar
1950 The analysis of linguistic borrowing. *Language* 26(2). 210–231.
Heath, Jeffrey
1978 *Linguistic diffusion in Arnhem Land*. Canberra: AIAS.
Heine, Berndt & Tania Kuteva
2005 *Language contact and grammatical change*. Cambridge: Cambridge University Press.
Hjelmslev, Louis
1938 Etude sur la notion de parenté linguistique: Relations de parenté des langues creoles. *Revue des Etudes Indo-Européenes* 2. 271–286.
Jespersen, Otto
1922 *Language: Its nature, development and origin*. London: Allen & Unwin.
Kihm, Alain
2003 Inflectional categories in creole languages. In Ingo Plag (ed.), *Phonology and morphology of creole languages*, 333–364. Tübingen: Max Niemeyer Verlag.
King, Ruth
2000 *The lexical basis of grammatical borrowing: A Prince Edward Island French case study*. Amsterdam: John Benjamins.

Kouwenberg, Silvia
1994 *A grammar of Dutch Berbice Creole and Cape Verdean*. New York: Mouton de Gruyter.
Kutscher, Sylvia
2001 *Nomen and nominales Synatgma im Lasischen. Eine deskriptive Analyse des Dialekts von Ardeşen*. München: Lincom Europa.
Law, Danny
2013 Inherited similarity and contact-induced change in Mayan Languages. *Journal of Language Contact* 6(2). 271–299.
Lefebvre, Claire
2004 *Issues in the study of pidgin and creole languages*. Amsterdam: John Benjamins.
List, Johann-Mattis, Shujulal Nelson-Sathi, Hans Geisler & William Martin
2013 Networks of lexical borrowing and lateral gene transfer in language and genome evolution. *Bioessays* 36. 141–150.
Matras, Yaron
2003 Mixed languages: Re-examining the structural prototype. In Yaron Matras & Peter Bakker (eds.), *The mixed language debate: Theoretical and empirical advances*, 151–176. Berlin: Mouton de Gruyter.
Matras, Yaron & Jeanette Sakel (eds.)
2007 *Grammatical borrowing in cross-linguistic perspective*. Berlin: Mouton de Gruyter.
McWhorter, John
1998 Identifying the creole prototype: Vindicating a typological claim. *Language* 74(4). 788–818.
McWhorter, John
2005 *Defining creole*. Oxford: Oxford University Press.
Meakins, Felicity
2011 Borrowing contextual inflection: Evidence from northern Australia. *Morphology* 21(1). 57–87.
Meakins, Felicity
2013 Mixed languages. In Yaron Matras & Peter Bakker (eds.), *Contact languages: A comprehensive guide*, 159–228. Berlin: Mouton.
Meakins, Felicity
2014 Nominals as adjuncts or arguments: Further evidence from language mixing. In Rob Pensalfini, Myfany Turpin & Diana Guilleman (eds.), *Language description informed by theory*, 283–315. Amsterdam: John Benjamins.
Meakins, Felicity
2015 From absolutely optional to only nominally ergative: The life cycle of the Gurindji Kriol ergative suffix. In Francesco Gardani, Peter Arkadiev & Nino Amiridze (eds.), *Borrowed morphology*, 189–218. Berlin: Mouton de Gruyter.
Meakins, Felicity, Xia Hua, Cassandra Algy & Lindell Bromham
2019 The birth of a new language does not favour simplification. *Language* 95(2).
Michaelis, Susanna, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.)
2013 *The survey of pidgin and creole languages vols I–IV.* Oxford: Oxford University Press.
Moravcsik, Edith
1978 Universals of language contact. In Joseph Greenberg (ed.), *Universals of human language: Vol. 1 Method and theory*, 95–122. Stanford CA: Stanford University Press.
Mufwene, Salikoko
1997 Kitúba. In Sarah G. Thomason (ed.), *Contact languages: A wider perspective*, 173–208. Amsterdam: John Benjamins.
Munro, Jennifer
2004 *Substrate language influence in Kriol: The application of transfer constraints to language contact in northern Australia.* Armidale: University of New England, PhD thesis.

Muysken, Pieter
1981 Halfway between Quechua and Spanish: The case for relexification. In Arnold Highfield & Albert Valdman (eds.), *Historicity and variation in creole studies*, 52–78. Ann Arbor: Karoma.
Muysken, Pieter
2000 *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.
Myers-Scotton, Carol
1993 *Duelling languages: Grammatical structure in codeswitching*. Oxford: Clarendon Press.
Myers-Scotton, Carol
2002 *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
Myers-Scotton, Carol
2003 What lies beneath: Split (mixed) languages as contact phenomena. In Yaron Matras & Peter Bakker (eds.), *The mixed language debate: Theoretical and empirical advances*, 73–106. Berlin: Mouton de Gruyter.
Nordhoff, Sebastian (ed.)
2012 *The genesis of Sri Lanka Malay: A case of extreme language contact*. Leiden: Brill.
O'Shannessy, Carmel
2013 The role of multiple sources in the formation of an innovative auxiliary category in Light Warlpiri, a new Australian mixed language. *Language* 89(2). 328–353.
Pakendorf, Brigitte
2009 Intensive contact and the copying of paradigms: An Ėven dialect in contact with Sakha (Yakut). *Journal of Language Contact* 2(2). 85–110.
Parkvall, Mikael
2008 The simplicity of creoles in cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, 265–285. Amsterdam: John Benjamins.
Plag, Ingo
2003a *Phonology and morphology of creole languages*. Tübingen: Max Niemeyer Verlag.
Plag, Ingo
2003b *Yearbook of morphology 2002*. London: Kluwer.
Roberts, Sarah & Joan Bresnan
2008 Retained inflectional morphology in pidgins: A typological study. *Linguistic Typology* 12(2). 269–302.
Ross, Malcolm
2006 Metatypy. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 95–99. Oxford: Elsevier.
Sakel, Jeanette
2007 Types of loans: Matter and pattern. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 15–29. Berlin: Mouton.
Sebba, Mark
1998 A congruence approach to the syntax of code-switching. *International Journal of Bilingualism* 2(1). 1–20.
Seifart, Frank
2012 The principle of morphosyntactic subsystem integrity in language contact: Evidence from morphological borrowing in Resígaro (Arawakan). *Diachronica* 29(4). 471–504.
Seuren, Pieter
1998 *Western linguistics: An historical introduction*. Oxford: Blackwell.
Singh, Rajendra
1982 On some 'redundant compounds' in Modern Hindi. *Lingua* 56(3–4). 345–351.

Smith, Ian, Scott Paauw & Bachamiya Hussainmiya
2004 Sri Lankan Malay: The state of the art. In Rajendra Singh (ed.). *Yearbook of South Asian languages 2004*, 197–215. Berlin: Mouton de Gruyter.
Thomason, Sarah G.
2015 When is the diffusion of inflectional morphology not dispreferred? In Francesco Gardani, Peter Arkadiev & Nino Amiridze (eds.), *Borrowed morphology*, 27–46. Berlin: De Gruyter Mouton.
Thomason, Sarah & Terrence Kaufman
1988 *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
Treffers-Daller, Jeanine
1994 *Mixing two languages: French-Dutch contact in a comparative perspective*. Berlin: Mouton de Gruyter.
Weinreich, Uriel
1974 [1953] *Languages in contact: Findings and problems*. The Hague: Mouton.
Whinnom, Keith
1971 Linguistic hybridization and the 'special case' of pidgins and creoles. In Dell Hymes (ed.), *Pidginization and creolozation of languages*, 91–115. Cambridge: Cambridge University Press.
Whitney, William
1881 On mixture in language. *TAPA* 12, 5–26.
Winford, Donald
2003 *An introduction to contact linguistics*. Malden, USA: Blackwell.

Felicity Meakins, Brisbane (Australia)

7. Orthography and graphemics

1. Conceptual preliminaries
2. Language contact in the written mode
3. Writing reforms
4. Orthographic choices in contact settings
5. Conclusion
6. References

1. Conceptual preliminaries

Language contact has been the object of extensive research over the last few years. Yet the analysis of oral data has prevailed to date, and as far as modern languages are concerned, only a relatively small body of studies have been devoted to language contact in the written medium. Written language entails the selection of a writing system, the choice of a script, and the development of an orthography. Because of their visual prominence and their association with religion, ethnicity, and cultural tradition, writing systems, scripts, and orthographies are salient markers of identity and are rarely, if ever, perceived as neutral instruments for recording speech.

In this chapter, the term *writing system* indicates a potential, abstract system of recording language, classified on the basis of its operational units, e.g. a phonetic-phonological

https://doi.org/10.1515/9783110435351-007

vs. a syllabic vs. a word-writing system; the actual graphic form of a writing system will be referred to as a *script*. Thus, Russian and English are written using similar writing systems whose basic functional units represent phonemes, yet the former employs the Cyrillic script, whereas the latter uses the Roman. The term *orthography* will be reserved for the language-specific set of conventions whereby a language is represented: different orthographies may exist for the same language, as in the case of British and American English, or in the case of German as used in Switzerland and in Germany. Note that, as pointed out by Coulmas (2013: 18), orthography and spelling are not synonymous, since "spelling, unlike orthography, can be deviant, that is, irregular or unusual", either purposely or not. In the presence of an established orthography any deviation from the norm has the potential to convey indexical meaning (Jaffe 2000; Jaffe et al. 2012; Sebba 2007, 2009).

For a long period, writing systems and orthographies were not regarded as suitable for study from a specifically sociolinguistic perspective. A pioneering attempt was the collection edited by Fishman (1977), which encompassed case studies documenting a variety of sociolinguistic settings. In the following decade, the idea that literacy is shaped by the social and cultural peculiarities of the community in which the reading and writing skills are deployed led to the development of the notion of literacy *as a social practice* (Barton 2006; Street 1984), within an approach known as New Literacy Studies. Specialized journals in various branches of sociolinguistics devoted single issues to the investigation of the non-linguistic questions affecting writing system, script choice, and spelling reforms. For example, in 1983, a special issue of the *International Journal of the Sociology of Language* devoted to literacy and ethnicity stressed the fact that writing could no longer be regarded as "a technology taught out of context" (Wagner 1983: 7), but entailed the consideration of religious, social, and practical factors. The analysis of the latter was also recommended by researchers focusing on development of orthographies for hitherto unwritten languages or in minority communities (Lüpke 2011; Roberts 2008; Unseth 2008). Lack of space prevents us from offering a detailed account of the publications flourishing in these fields (see Blommaert [2013], Bunčić [2017: 27–50], and Sebba [2009] for in-depth overviews). In the next section, some of the key methodological and theoretical notions that guide our understanding of the dynamics of language contact in the written mode will be examined and discussed.

2. Language contact in the written mode

Traditional approaches to contact linguistics point to bilingual speakers as the agents of contact-induced change (Matras 2009). In a similar way, those responsible for *written* language contact are bi- or multiliterate, i.e. literate in two or more scripts, though, predictably, with differing degrees of competence in and acquaintance with written materials. Historically, most scripts and orthographies were devised by bi- or multiliterate writers who adapted and extended written models to their particular context. As Sebba (2007: 162) illustrates, script choice and orthography development are mediated by bilingual literate elites who devise a set of conventions under the influence of the rules already familiar to them for writing another language:

> […] languages and their associated orthographies do not exist in isolation. Their users know something − a little or a lot − about other languages or language varieties which use orthographies different from theirs, and are able to make use of this knowledge in various ways.

The most common way to "make use of this knowledge" is to express affiliation with (or distance from) a certain religious, ethnic, or linguistic group by selecting a script, or by choosing/rejecting graphemes, character combinations, and diacritics that are perceived as emblematic of the group in question.

Sebba (2015: 209) illustrates "three processes" whereby social meaning can be ascribed to scripts and spelling choices. The first one, *attribution*, is the process whereby "one group of people, A, make an association between a linguistic feature or language-related practice, X, and a group of people, B, who (supposedly) use that feature or engage in that practice". Mumin (2014: 48), for example, reports that in Sub-Saharan Africa, all texts written in Ajami − e.g. Swahili, Wolof, or Fulfulde poems and tales − tend to be perceived by non-Muslims as Islamic literature, irrespective of their contents. In other words, the choice of Ajami, an adapted version of the Arabic *abjad* (from Arabic *ʿağam* 'non-Arab, Persian', since Persian was the first language other than Arabic to be written in Arabic script, see Mumin [2014: 41]), becomes associated with − or more precisely, is attributed to − Muslim religion.

Attribution is a prerequisite for *iconisation*, a process originally described in the following way by Irvine and Gal (2000: 37): once a linguistic feature has become associated with a certain social group, the same feature "appear[s] to be an iconic representation of [the group itself], as if a linguistic feature somehow depicted or displayed a social group's inherent nature or essence". To exemplify this process, Sebba (2015: 116) recalls the account offered by Ahmad about the line of thinking which led to the stigmatization of the Urdu script − and hence, of the Muslim minority − in late nineteenth century India: "it was argued that the fraudulent character of the script and language stemmed from their orthographic shortcomings […]. This 'promiscuous' nature of the script was then held responsible for causing fraud and forgery".

A third process, which Sebba (2015: 19) calls *branding*, entails "the attribution of a graphical element − such as a letter of the alphabet or a diacritic mark − to a group, with or without iconisation". By definition, branding entails the selection of one (or more) element(s) from the inventory of graphemes and diacritics of a certain orthographic tradition. This may be done "by the users themselves, who establish the item in question as their 'brand', or it may be done by an outgroup" (Sebba 2015: 213). A case in point is the rejection of the orthography for Cimbrian, a Germanic (Bavarian) minority language spoken by roughly 300 speakers in the Northern Italian village of Lucerna. As recounted by Bidese (2015: 129), the most controversial issue was the decision to represent the voiced fricative /ʒ/ by using <z>, instead of <s>. The choice was supposed to break with both the Italian and the German orthographic traditions and set Cimbrian apart from other German minorities in Northern Italy. Yet, the local community perceived the presence of <z> as disturbing, since it would prevent their language from 'looking like' German. Thus, a single visual element, <s>, was established as a 'brand' marking the Germanness of the Cimbrian community and its language.

In the processes just described, social meaning "is mediated by the role of script-as-image" (Sebba 2015: 209), i.e. by the tendency to associate the visual salience of a

certain graphical feature with specific religious, cultural, and/or social groups. In Northern Nigeria, Ajami was replaced by the Roman script at the beginning of the twentieth century. The latter, locally called Boko, was increasingly seen as an iconic representation of Western colonialism, a process of attribution recently exploited by the Boko Haram terrorist movement, whose name is generally translated as 'Western education is forbidden', thus equating the Roman script with Western culture and education (Warren-Rothlin 2014: 265).

In most parts of the world it is possible to identify "a default writing system", the unmarked choice when the orthography for a previously unwritten language is developed (Coulmas 2014b: 3). Generally, the default choice coincides with the writing system taught within the local educational system and/or with the one used in order to write the language(s) enjoying statutory recognition. Since the latter, by virtue of their status and literary tradition, enjoy higher prestige than varieties lacking a written form, the choice tends to fall on the writing system of the literary language(s) when vernacular writing is established. Hence, most African languages are presently written in an adapted version of the Roman script (Pasch 2008), the script in use for writing the major former colonial languages. The choice is often justified in terms of transferability, i.e. as a practical benefit minimizing the effort associated with the development of literacy skills in different languages. And indeed, in most developing countries, literacy in a community language can be extended to a language of wider communication only if the same script is used. However, as many African intellectuals have denounced (Bamgbose 2000; Batibo 2005; Djité 2008), the prominence of the Roman alphabet in computer-mediated communication, and the interplay of political and economic forces favoring Western cultural practices over local ones, also play a major role.

In Medieval Europe, the question of whether or not the Roman script should be adopted to write the various Romance vernaculars never really came up for debate. Latin, the written language *par excellence*, was taken as a model for the orthography of all emerging literary languages, including non-Romance languages like English and German, which had independently developed their own runic scripts. The demise of runic writing was encouraged by the Catholic Church, whose representatives iconified runes as associated with the Germanic pagan faith and the magical rituals performed by its adherents (Antonsen 2002: 37−39). The inventory of graphemes of the Roman alphabet, however, was tailored to the Latin phonological system, and thus inadequate for the needs of some Germanic languages. Two traditional Old English graphemes <þ> *thorn* and <ð> *eth*, for instance, were replaced by the digraph <th>, which fails to convey the voiced /ð/ vs. voiceless /θ/ distinction. Another case in point is the underrepresentation of the English vowel system resulting from the Latin alphabet's paucity of vowel letters (Coulmas 2003: 89−108).

This leads us to one of the most common outcomes of written language contact, i.e. *digraphia* (Dale 1980; DeFrancis 1984; Grivelet 2001). A sociolinguistic typology of societies employing two or more scripts, writing systems, or orthographies for the same language has recently been devised by Bunčić, Lippert, and Rabus (2016), who adopted *biscriptality* as a cover term encompassing nine basic types of language situations that emerge from the intersection of two heuristic categories: the sociolinguistic factors influencing the choice of a script/writing system, and the graphematic levels affected by variation. In situations of competing scripts and orthographies, the selection of a writing

system or a given orthographic variant acquires great symbolic value, and may be interpreted as an act of identity, comparable to language choice in bi- or multilingual speech.

In Medieval Spain, Arabic used to be written in Arabic script (by Muslims), in Roman script (by Christians) and in Hebrew script (by Jews), the choice representing a powerful symbol of group membership which indexed the writer's affiliation with a certain religious and/or cultural tradition (Warren-Rothlin 2014: 264).

When the Kingdom of Serbs, Croats, and Slovenes was established at the end of the First World War, Serbo-Croatian was written in the Cyrillic alphabet by the Orthodox Serbs, and in the Roman script by the Catholic Croats. The totalitarian fascist regime that controlled the region from 1941 to 1945 banned the use of the Cyrillic script not only in public, but also in private correspondence. As a reaction against these restrictions, Tito's socialist Yugoslavia granted equal status to both the Roman and the Cyrillic alphabets, and issued a common orthography for Serbo-Croatian. When Yugoslavia collapsed in the 1990s and the Serb and Croat leaders proclaimed their own self-determining republics, the return to different scripts contributed to the creation of a clearer boundary between the two newly-established countries, whose political legitimacy could not yet be taken for granted (Bunčić, Lippert, and Rabus 2016: 177−180).

One of the lessons from this case is that there appears to be a correlation between the successful implementation of writing reforms and periods of socio-political unrest, when new allegiances are sealed, everyday certainties collapse, and people are more inclined to accept change. But the relation is in fact bi-univocal: as argued by Fishman (2006: 4), language reforms are sometimes an outcome of socio-political change, and sometimes a tool for bringing socio-political change into being. Why this may be so is the subject of the next section.

3. Writing reforms

As suggested by Coulmas (2003: 234−235), we distinguish three kinds of reform: orthographic reforms, involving the adjustment of spelling conventions; script reforms, entailing the replacement of one graphical code with another (e.g. a shift from the Cyrillic to the Roman alphabet, as in the case of the Azeri language illustrated below); and writing system reforms, concerning the overall design of the writing system and its inventory of analytical units, as well as the relationship between graphic and linguistic units. The latter are obviously the most consequential of the three, given their impact on the reading and writing skills of literate language users.

Writing reforms may be justified in strictly linguistic terms by the need to reduce the sign inventory, simplify spelling conventions, and make them more functional and transparent. Improvements are of course always possible, but most commonly, linguistic concerns conceal other motivations. The resistance with which script and spelling reforms are met − however minimal the changes proposed − bears witness to the symbolic significance that people ascribe to the written form of their language(s) (Coulmas 2013; Johnson 2005; Sebba 2007). This is all the more so in language contact situations, where the possibilities of socially-meaningful variation multiply.

3.1. Cultural factors

The pressure to move towards Western culture and to partake in the social and economic benefits associated with life in a Western country is one of the major arguments in favor of a shift to the Roman script. Coulmas (2014a: 296) recounts how, within the Chinese cultural sphere, three languages had traditionally relied on the Chinese writing system: Vietnamese, Korean, and Japanese. While the agglutinative morphology and polysyllabic word structure of Korean and Japanese make Chinese characters highly inadequate to represent these languages, Vietnamese is typologically similar to Chinese and can be effectively written with Chinese characters. Yet at the beginning of the twentieth century, Vietnamese reformers opted for a Romanized orthography, originally introduced by French Jesuit missionaries, in an attempt to modernize their country and resist Chinese cultural and economic domination.

In northern Nigeria and Cameroon, centuries-old traditions of Ajami writing were replaced by Roman script writing on the initiative of Christian missionaries, not as an anti-Islamic measure, but in order to ease the dissemination of Western culture and values (Warren-Rothlin 2014: 262). This policy was so successful that official literacy statistics even today only account for Roman-script literacy, disregarding the languishing competence in Ajami.

3.2. Political factors

The reforms implemented by Turkey's first president, Mustafa Kemal Atatürk, as part of a modernization strategy in the newly proclaimed Republic of Turkey, included the introduction of an adapted version of the Roman script for writing Turkish, which had been written in Arabic script for centuries under Ottoman rule. The version of the Roman alphabet adopted for that purpose was shrewdly referred to by the name 'Turkish letters'. The use of two modified graphemes (<ı> for /ɯ/ and <ş> for /ʃ/) that were swiftly established as a brand of Turkish authenticity (Coulmas 2016: 208) favored the acceptance of the new script.

Azerbaijan is another case in point. Since the seventh century, the Azeri language had been written in Arabic script. After the Soviet Union annexed Azerbaijan in 1920, the Soviets encouraged Azeri literacy in Roman script as a move to sever the link with Iran, a powerful Muslim neighbor and home to a large Azeri-speaking population. The unexpected shift to Roman script implemented by Turkey in the same years, however, increased concerns about the development of a pan-Turkic movement in the Muslim-majority republics of the Soviet Union. Hence, in 1939, Stalin declared that the Cyrillic script would replace the Roman alphabet for all the languages of the USSR (Hatcher 2008: 109). In 1991, when the Soviet Union collapsed and many of its former component republics − including Azerbaijan − began to seek political independence, one of the first steps taken by the Azerbaijani parliament was to change the script back to Roman, thus asserting the country's cultural autonomy from Moscow.

Hong Kong and Taiwan's rejection of character simplification schemes promulgated by the People's Republic of China is another example of how political arguments may influence the attitudes of a speech community towards writing reforms.

3.3. Religious factors

The influence of religion on script and writing reforms is self-evident: the major Abrahamic religious traditions − Judaism, Christianity, and Islam − associate access to the divine with access to sacred texts, whose visual appearance is viewed as eternal and immutable as the truths that the same texts disclose. Besides, as Sebba observes, "a script provides a symbolic link with others who share the same religion or supposed cultural origins, even where they do *not* [emphasis mine] share the same language" (Sebba 2007: 142). Thus, Maltese, an Afro-Asiatic language of the Semitic branch, typologically similar to Arabic, is written in Roman script as a result of the long-established religious prominence of Roman Catholicism in Malta.

One of the major obstacles faced by the script reform promoted by Mustafa Kemal Atatürk was the opposition of the local Islamic clergy: "in 1910, [Turkey's] religious authorities issued a fatwa to the effect that, for a Muslim, using a script other than Arabic was a sin" (Coulmas 2016: 207). The reformers succeeded in winning the people's resistance by underscoring the advantages of the new script (e.g. a larger inventory of vowel symbols), without calling into question their religious convictions.

Tuareg, a Berber language spoken in Algeria, Morocco, Mali, and Niger, was originally written in Tifinaġ, an autochthonous alphabet derived from the ancient Numidiam script (Pasch 2008: 72). From the seventh century onward, as a result of Arab expansion in North Africa, this script was gradually replaced by the Arabic script until Tifinaġ writing was limited almost exclusively to domestic usage. In the late nineteenth century, French colonization brought the French language, written in Roman script. For decades, three scripts were used in order to write Tuareg in Morocco, the representatives of the Tuareg minority being unable to agree on a shared solution. Recently, despite the functional merits and the spread of both Arabic- and Roman-script literacy, the choice has fallen on Tifinaġ (Savage 2008), whose grapheme *yaz*, incidentally, is prominently featured on the Berber flag (Lüpke 2011: 323).

As the above examples show, cultural, political, and religious factors are often intimately intertwined. As for the correlation between writing reforms and periods of sociopolitical unrest, note that most reforms tend to be top-down measures imposed upon language users by nationalistic, authoritarian governments: an effective control over the territory and the population concerned is often necessary to minimize the resistance of those literate in the old script and to contain the discontent arising out of the quandaries associated with the change.

4. Orthographic choices in contact settings

4.1. The iconic value of single graphemes

Orthographic reforms are admittedly more frequent and less consequential than script and writing system reforms, yet the identity-marking potential of orthographic choices makes them no easier to implement. Orthography is one of the most visible features of a written language, hence even a small change, such as one involving the introduction of diacritics or a certain combination of graphemes, may have identity-related implications.

When political authorities in the Cape Verdean archipelago − a group of tiny islands situated in the North Atlantic off the coast of Senegal − inaugurated the policy of *bilinguismo assumido* ('conscious bilingualism'), involving co-official status for Portuguese and the local Portuguese-based creole, the need to provide the latter with a unified orthography became urgent. In 1998, a Roman-script orthography (called ALUPEC, an acronym for *Alfabeto Unificado para a Escrita do Cabo-verdiano*) was officially introduced. As in most post-colonial contexts, the ruling principle of ALUPEC was to underscore the distance between the creole and its lexifier language (Canù 2015: 137). To this end, orthography planners decided that ALUPEC would make use of <k> for the expression of the voiceless velar plosive [k], instead of <c> and <q>, which represent the same velar sound in Portuguese orthography. To be sure, the visual prominence of this single graphic element, along with the high frequency of occurrence of [k], enables the reader to establish at first sight whether a certain text is creole or Portuguese. Analysis of the local linguistic landscape (Canù 2015), however, reveals that, nearly twenty years after the introduction of ALUPEC, the major preoccupation of creole writers seems to be the graphical rendition of the phonetic peculiarities of the main vernacular varieties spoken in the archipelago, rather than the expression of a creole identity as opposed to a Portuguese one.

4.2. Orthographic assimilation of loanwords

It is intuitively evident that the adoption of lexical borrowings from a donor into a recipient language may be favored by the presence of the same script. In this case, lexical borrowings may either retain their original orthographic shape (e.g. the English loanwords *football*, *leader*, and *smoking* ['suit'] in present-day Italian), or display various degrees of orthographic adaptation to the recipient language (e.g. *fùtbol*, *líder*, and *esmoquin* in present-day Spanish). The choice largely depends on language-specific preferences, and does not necessarily correlate with consistent patterns of phonological adaptation.

In language contact situations also involving script contact, the possibilities multiply. As suggested by Klöter (2010), loanwords analysis entails three concurrent layers: i) form (signifier), ii) meaning (signified), and iii) graph, i.e. the graphic representation of a linguistic unit. The outcomes of written language contact may thus be represented by a formula in which each layer is preceded either by a plus sign (if it is retained from the donor language) or by a minus sign (if it is not). We will leave aside the patterns that are less relevant for the present analysis − e.g. the occurrence of the English words *file* and *skateboard* in a Russian newspaper article otherwise written in Cyrillic script, which Klöter (2010) would describe as [+form, +meaning, +graph] loanwords, but which other scholars consider virtually indistinguishable from written single-word switchings (Haspelmath 2009: 40). We instead illustrate this point with a couple of examples.

The linguistic landscape of Asmara, Eritrea's capital city, is dominated by the presence of three languages: Tigrinya, Eritrea's national language (written in Ethiopic script, an adapted version of the Geez syllabary), English, and Italian, the latter being a former colonial language alongside English. In Asmara city center, the window of an ice cream shop may exhibit the text ላ ጀላተርያ /la ʤelaterija/ (Vergari 2015), a loanword from Italian, which retains the same meaning and almost the same phonological form of the

donor language *la gelateria* /la ʤelateˈria/ ‘the ice cream shop’, but is written in Ethiopic script, and hence, hardly recognizable as a word of foreign origins. The same holds for the restaurant name ፒሰሪያ ናፖሊ /piserija napoli/, or for a banner reading ኤለትሪቺስታ /ʔeletriʧista/ ‘electrician’ (from Italian *elettricista* /eletːriˈʧista/). Minor discrepancies between the phonological form of the loanwords in the donor and recipient languages are due to restrictions imposed by the use of the Geez syllabary, which lacks characters for the expression of the voiceless alveolar affricate /ʦ/ occurring in *pizzeria* /piʦːeˈria/ and for the notation of isolated vowel sounds like /a/. Those sounds are spelled as ሰ /se/ and ያ /ja/, respectively. For the same reason, word-initial /e/ in *elettricista* is spelled ኤ /ʔe/, and consonant length (which conveys phonological meaning in Italian) is not notated. In Klöter’s (2010) terms, these are instances of [+form, +meaning, −graph] loanwords.

Loans from Russian into Italian, on the other hand, are transliterated into the Roman alphabet, but generally undergo phonological assimilation: hence *перестройка* (‘political reform’) is written *perestrojka* and pronounced /pereˈstrɔjka/. In a similar way, Warren-Rothlin (2014: 264) reports the use of Ajami on the labels of Western foodstuffs in Northern Africa, e.g. كوكا كولا < kwkʾ kwlʾ > ‘Coca Cola’. Both examples fit Klöter’s (2010) description of [−form, +meaning, −graph] loanwords.

4.3. Orthographic choices in networked multilingualism

The digital revolution has brought a complete change in the technology of the dissemination of information, and has paved the way for previously unknown vernacular literacy practices. *Networked multilingualism* is a cover term recently proposed (Androutsopoulos 2015) for the theorizing of the linguistic practices shaped by three sets of constraints: mediation of written language by digital technologies, access to network resources, and orientation to networked audiences. The analysis of the orthographic choices emerging in digital arenas, such as internet fora and blogs, Facebook and text-messaging provides rich material for the study of multilingualism and written language contact as modern phenomena (Deumert 2014; McLaughlin 2014).

4.4. Typographic mimicry

Another practice − also favored by the recent emergence of digital writing tools − concerns the recourse to *faux fonts* (Sebba 2015: 219), or typographic mimicry, i.e. “the identification and use of certain graphic features to mark an inscription as foreign, while ensuring the transmission of the contents” (Coulmas 2014b: 17). Consider, for instance, the following text, displayed on the shop sign of a store selling imported food products in Amsterdam: هعوق یعالة (Dutch *alles goed?*). Another case in point comes from an article in *Time Magazine* (25 March 2013, page 33) that discusses the possible impact of the Arab Spring protests, and is accompanied by the following title, ےلم ےہم ےلد۲س ‘we the people’, designed to resemble Arabic script. These and similar examples suggest that the distinctive features of a certain writing system may be identifiable even by a viewer who is not literate in the writing system in question, and who may extract from its sign

inventory a set of visual peculiarities (e.g. the prevalence of diacritics above or below the letters, the characteristic shape of graphemes, and the use of certain digraphs or character combinations) to establish those peculiarities as a brand of a group to whom the writing system is attributed. This leads us back to the crucial role played by certain graphic elements, whose visual salience can be exploited in order to index individual users or groups.

5. Conclusion

Language contact in the written mode relies on the visual channel. Written language is not merely a means of representing speech, but entails processing images, which are seen and remembered: "although they do not directly carry semantic meaning, these images may nevertheless provide contextualization for the message they compose […] and may in addition carry social meaning" (Sebba 2015: 208).

The technology of writing encompasses both instrumental and symbolic functions (Coulmas 2013), and when it comes to implementing orthography or script reforms, the latter outweigh the former. Language users display an unexpected degree of awareness of the symbolic significance of script and spelling choices, and exploit the visual salience of some graphic elements in order to express affiliation to (or disaffiliation from) certain religious, ethnic, and/or linguistic groups. A thorough investigation of the dynamics of written language contact will contribute to a better understanding of the privileged sites for the negotiation of individual and group identity.

6. References

Ahmad, Rizwan
2012 Hindi is perfect, Urdu is messy: The discourse of delegitimation of Urdu in India. In Alexandra Jaffe, Jannis Androutsopoulos, Mark Sebba & Sally Johnson (eds.), *Orthography as Social Action. Scripts, spelling, identity and power*, 103−133. Berlin: De Gruyter.

Androutsopoulos, Jannis
2015 Networked multilingualism. Some language practices on Facebook and their implications. *International Journal of Bilingualism* 19(2). 185−205.

Antonsen, Elmer
2002 *Runes and Germanic linguistics*. Berlin: De Gruyter.

Bamgbose, Ayo
2000 *Language and exclusion: The consequences of language policies in Africa*. Hamburg: LIT.

Barton, David
2006 *Literacy: An introduction to the ecology of written language*. Oxford: Blackwell.

Batibo, Herman
2005 *Language decline and death in Africa.* Clevedon: Multilingual Matters.

Bidese, Ermenegildo
2015 La guerra della <s> e della <z> a Luserna. In Silvia Dal Negro, Federica Guerini & Gabriele Iannàccaro (eds.), *Elaborazione ortografica delle varietà non standard*, 119−134. Bergamo: Bergamo University Press.

Blommaert, Jan
2013 Writing as a sociolinguistic subject. *Journal of Sociolinguistics* 17(4). 440–459.
Bunčić, Daniel
2017 Factors influencing the success and failure of writing reforms. *Studi Slavistici* 14 (special issue: 1917–2017: One hundred years of graphic (r)evolution in the Soviet Space). 21–46.
Bunčić Daniel, Sandra L. Lippert & Achim Rabus (eds.)
2016 *Biscriptality: A sociolinguistic typology*. Heidelberg: Winter.
Canù, Silvia
2015 Tendenze nelle scelte dei parlanti nelle grafie spontanee in Creolo capoverdiano. In Dal Negro Silvia, Federica Guerini & Gabriele Iannàccaro (eds.), *Elaborazione ortografica delle varietà non standard*, 135–157. Bergamo: Bergamo University Press.
Coulmas, Florian
2003 *Writing systems: An introduction to their linguistic analysis*. Cambridge: Cambridge University Press.
Coulmas, Florian
2013 *Writing and society: An introduction.* Cambridge: Cambridge University Press.
Coulmas, Florian
2014a Authentic writing. In Véronique Lacoste, Jakob Leimgruber & Thiemo Breyer (eds.), *Indexing authenticity: Sociolinguistic perspectives*. 289–303. Berlin: De Gruyter.
Coulmas, Florian
2014b Writing systems and language contact in the Euro- and Sinocentric worlds. *Applied Linguistic Review* 5(1). 1–21.
Coulmas, Florian
2016 *Guardians of language: Twenty voices through history*. Oxford: Oxford University Press.
Dale, Ian
1980 Digraphia. *International Journal of the Sociology of Language* 26. 5–13.
DeFrancis, John
1984 Digraphia. *Word* 35. 59–66.
Deumert, Ana
2014 *Sociolinguistics and mobile communication*. Edinburgh: Edinburgh University Press.
Djité, Paulin
2008 *The sociolinguistics of development in Africa*. Clevedon: Multilingual Matters.
Fishman, Joshua (ed.)
1977 *Advances in the creation and revision of writing systems*. The Hague: De Gruyter.
Fishman, Joshua
2006 *DO NOT leave your language alone*. Mahwah, NJ & London: LEA.
Grivelet, Stéphane (ed.)
2001 Digraphia: Writing systems and society. [Special issue]. *International Journal of the Sociology of Language* 150.
Haspelmath, Martin
2009 Lexical borrowings: Concepts and issues. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 35–54. Berlin: De Gruyter.
Hatcher, Lynley
2008 Script change in Azerbaijan: Acts of identity. *International Journal of the Sociology of Language* 192. 105–116.
Irvine, Judith & Susan Gal
2000 Language ideology and linguistic differentiation. In Paul V. Kroskrity (ed.), *Regimes of language: Ideologies, polities and identities*, 35–83. Oxford: James Currey.

Klöter, Henning
2010 What is being borrowed? Language and script contact in Taiwan. In Alex de Voogt & Irving L. Finkel (eds.), *The idea of writing: Play and complexity*, 93–115. Leiden & Boston: Brill.
Jaffe, Alexandra
2000 Non-standard orthography and non-standard speech. *Journal of Sociolinguistic* 4(4). 497–513.
Jaffe, Alexandra, Jannis Androutsopoulos, Mark Sebba & Sally Johnson (eds.)
2012 *Orthography as social action: Scripts, spelling, identity and power*. Berlin: De Gruyter.
Johnson, Sally
2005 *Spelling trouble? Language, ideology and the reform of German orthography*. Clevedon: Multilingual Matters.
Lüpke, Friederike
2011 Orthography development. In Peter Austin & Julia Sallabank (eds.), *The Cambridge handbook of endangered languages*, 312–336. Cambridge: Cambridge University Press.
Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.
McLaughlin, Fiona
2014 Senegalese digital repertoire in superdiversity: A case study from Seneweb. *Discourse, Context and Media* 4(5). 29–37.
Mumin, Meikal
2014 The Arabic script in Africa: Understudied literacy. In Meikal Mumin & Kees Versteegh (eds.), *The Arabic script in Africa*, 41–76. Leiden & Boston: Brill.
Pasch, Helma
2008 Competing scripts: The introduction of the Roman script in Africa. *International Journal of the Sociology of Language* 191. 65–109.
Roberts, David
2008 Thirty years of orthography testing in West African languages (1977–2007). *Journal of West African Languages* XXXV(1–2). 199–242.
Savage, Andrew
2008 Writing Tuareg: The three script options. *International Journal of the Sociology of Language* 192. 5–13.
Sebba, Mark
2007 *Spelling and society: The culture and politics of orthography around the world.* Cambridge: Cambridge University Press.
Sebba, Mark
2009 Sociolinguistic approaches to writing system research. *Writing System Research* 1(1). 35–49.
Sebba, Mark
2015 Iconisation, attribution and branding in orthography. *Written Language and Literacy* 18(2). 208–227.
Street, Brian V.
1984 *Literacy in theory and practice*. Cambridge: Cambridge University Press.
Unseth, Peter (ed.)
2008 The sociolinguistics of script choice. [Special issue]. *International Journal of the Sociology of Language* 192.
Vergari, Moreno
2015 Italian in the linguistic landscape of Asmara (Eritrea). *Ethnorema* 11. 95–106.
Wagner, Daniel A.
1983 Ethno-graphies: An introduction. *International Journal of the Sociology of Language* 42. 5–8.

Warren-Rothlin, Andy
2014 West African scripts and Arabic-script orthographies in socio-political context. In Meikal Mumin & Kees Versteegh (eds.), *The Arabic script in Africa*, 261–289. Leiden & Boston: Brill.

Federica Guerini, Bergamo (Italy)

8. Levels of representation in phonetic and phonological contact

1. Introduction
2. Linguistic dominance and agentivity
3. The sound system
4. Contact and modularity
5. Conclusion
6. References

1. Introduction

For people who speak more than one language, any linguistic domain may be affected through contact. This chapter explores those influences in phonetics and phonology – the *sound system*. I concentrate on segmental representations, but prosodic and phonotactic considerations are also amenable to the analysis. At the heart of language contact is a bilingual's "rearrangement of patterns that result[s] from the introduction of foreign elements into the more highly structured domains of language" (Weinreich [1953] 1966: 1). The composition of these structured domains guides our understanding of contact patterns of bilingual individuals and communities.

Bilingual sound patterns manifest themselves in many ways, from a 'foreign' accent to loanword integration. Weinreich (1966: 11) defines two types of interference by distinguishing between the spontaneous speech of a bilingual and the forms that are adopted into the language independent of bilingualism. Additionally, Haugen (1950: 212) differentiates between substitutions and importations. The former occurs as an imperfect approximation of a pattern in one language based on the system of another; the latter describes a reproduction that is identical to the target language pattern. For example, the English loanword in Norwegian *hyppa* 'whip' substitutes Norwegian [h] and [y] for English [ʍ] and [ɪ], respectively; [h] is also an importation because it is indistinguishable in Norwegian and English (Haugen 1950: 213). These two analyses are not mutually exclusive. Haugen's substitutions and importations define speech interferences based on similarities and differences between the two language systems, whereas Weinreich's interferences characterize changes that occur as a result of bilingual forms that are integrated into the language, even among monolingual speakers.

Van Coetsem (1988) approaches Weinreich's speech interference and Haugen's substitutions in terms of transfer from the source language (SL) to the recipient language (RL). The sounds produced in the RL are adjusted to those of the SL when the speaker

https://doi.org/10.1515/9783110435351-008

is more proficient (linguistically dominant) in the source language (Section 2). For example, native French speakers often associate English aspirated stops with the French voiceless series, producing unaspirated French [p] for English [p^h] (van Coetsem 1988: 11). Such substitutions tend to operate on phonological categories. However, research in second language acquisition finds that L2 learners classify similar L1 and L2 sounds under the same phonetic category, and that increased acquisition of an L2 influences L1 production (e.g. Flege 1995). These and similar patterns of transfer are driven by the sound systems of the languages in contact. The outcomes are dependent on the extent to which bilinguals learn and use their languages' phonological and phonetic systems, and their ability to maintain those systems over the lifespan (Matras 2009: 225).

An account of phonological and phonetic contact must consider bilingualism both as an individual cognitive phenomenon and as a social group behavior. I focus here on the structural components of contacting sound systems, and the processes that underlie their interactions, by drawing on bilingual examples of VOICING distinctions (adopting Docherty's [1992] convention to signify the opposition in uppercase). Specifically, I discuss interactions between 'fortis-lenis' and 'voicing' systems, in which laryngeal contrasts are based on the features [spread (glottis)] in the former and [slack (vocal folds)] in the latter (Section 3). These examples illustrate the ways in which varying degrees of proficiency in the L2 produce asymmetric patterns of transfer in distinct parts of the sound system, e.g. the imposition of phonological categories, the introduction of new articulatory gestures, and granular adjustments to phonetic productions. The results of these processes introduce novel phonetic, allophonic, and phonological forms into a bilingual speech community and contribute to rich variation (e.g. Dubois and Horvath 1998; Purnell et al. 2005; Sharma 2011). I leave aside the social aspects of the allocation, adoption, and spread of these forms over time and space.

I discuss contact theory and the two transfer types in language contact in Section 2. This theory operationalizes linguistic dominance and the relative stability of linguistic systems to make predictions on the direction of transfer from one language to another. In Section 3, I introduce the architecture of the sound system adopted here. This position outlines different domains that are transfer targets. I analyze examples from four case studies in Section 4, and show that a modular theory of phonology adds nuance to the contact patterns, and predicts outcomes with respect to a bilingual's linguistic abilities, including changes in abilities over time. These examples support distinct types of phonological transfer based on the portion of the sound systems transferred in language contact. I present concluding remarks in Section 5.

2. Linguistic dominance and agentivity

When individuals speak more than one language, one is often (linguistically) dominant, i.e. the language in which a bilingual is most proficient regardless of the order of acquisition and primary community language (Winford 2005: 377). Uneven proficiencies produce asymmetric transfer of linguistic material across languages (Howell 1993: 189; Winford 2005). However, dominance relationships can, and do, change over a speaker's lifespan, and may result in attrition, or linguistic forms that diverge from previously acquired structures. Putnam, Perez-Cortes, and Sánchez (to appear) view these changes

as a consequence of differences in bilingual and monolingual speakers' activation of features for comprehension and production, followed by a reanalysis of the relationships between features and structures across languages. As a bilingual processes and activates features of a less-dominant language through more frequent and consistent use, dominance in that language increases to the extent that the speaker may achieve similar levels of proficiency in each language. Conversely, infrequent use of a dominant language weakens grammatical associations, and dominance may shift between two or more languages, as a bilingual whose L2 fluency surpasses that of their L1 (i.e. patterns associated with L1 attrition). Depending on the relative degrees of use, dominance shifts may occur in a number of independent linguistic domains, or as a general shift in fluency from one language to another. Varying levels of proficiencies and linguistic dominance have implications for both the type and degree of a language's influence on another, and whether these influences concern a part or the whole of the sound system (Sections 3 and 4). A more activated portion of the grammar may become dominant in one language relative to another, and therefore be prone to phonological impositions (Howell 1993; van Coetsem 1988).

To differentiate the interactions between dominant and non-dominant languages, van Coetsem (1988: 3) defines two transfer types − imposition and borrowing − according to whether the speaker's dominant language is the SL or the RL. Speakers are agents as either SL or RL speakers; agentivity in one of the two languages distinguishes the two transfer types. Imposition is the result of SL agentivity, where the dominant language is the source of linguistic material transferred to the less-dominant RL.

Borrowing is the result of RL agentivity. Here the less-dominant SL transfers material into the dominant RL (van Coetsem 1988: 3). For example, Norwegian immigrants in the United States adopted English words in their linguistically dominant Norwegian vocabularies, including *hyppa* 'whip' and *krækkis* 'crackers' (Haugen 1950: 222). Whether a language is the RL or the SL depends not on the dominance relationship between the languages, but on the direction of transfer between the two languages. Furthermore, the type of material prone to impositions and borrowings differs. Van Coetsem (1988: 25) refers to this asymmetry as the 'stability gradient', or as Winford explains, "speakers tend to preserve the more stable components of the language in which they are more proficient" (2005: 377). More stable linguistic domains (e.g. phonology and syntax) are more affected by SL agentivity, whereas the less stable domains (e.g. lexical items) are more open to borrowing into the dominant language through RL agentivity (Howell 1993: 189). Both imposition and borrowing may operate in conjunction, as in my own analysis of English loanwords in the Norwegian spoken in the United States (Natvig 2017). RL agentivity targets the less stable (lexical) domain of the non-dominant English, introducing 'whip' and 'crackers' into the Norwegian vocabulary. Their Norwegian pronunciations show the interpretation of English sounds through the dominant Norwegian phonological system, which imposes those categories on the loanwords through SL agentivity.

Phonology is a more resilient domain of the grammar (Benmamoun, Montrul, and Polinsky 2013: 141), yet evidence of L2 phonetic influence on L1 phonology (Section 4) suggests that stability is not uniform across the whole sound system. I next discuss how multiple levels of representation within the sound system models different types of transfer according to the relative stability of the phonology and the phonetics, and the degree of linguistic dominance in a given portion of the sound system.

3. The sound system

Van Coetsem (1988) considers a bilingual's linguistic dominance and the agency that underlies the transfer of linguistic material. The theory allows us to predict what material is susceptible for transfer based on linguistic dominance, and how changing language proficiencies interact with that process. Equally important for the latter is the structure of the phonological and phonetic components of the grammar. Consider three distinct levels of representation within the sound system: the Phonological level, the Phonetic-Phonological level, and the Phonetic level (Purnell and Raimy 2015). Each operates on different types of units, governs different processes, and are target domains for the transfer of different material (Section 4).

The Phonological level organizes a language's phonemes into categories based on how sounds differ from each other. These representations encode labels for the sounds that carry language-specific lexical and grammatical distinctions, or contrasts (Goldsmith 1995). While theories differ regarding the substance of features, it is generally accepted that the full spectrum of features in the speech signal is not required to mark phonological contrasts. For example, /o, u/ are the only back, round vowels in a language with the five vowel phonemes /i, e, u, o, a/. A language may designate /o, u/ as either [back] or [round]. The language need not use both [round] and [back] because it does not have non-round, back vowels nor round, front vowels. The features that mark a given phoneme are determined by the generalizations that learners make from a language's phonological alternations. Active features in those patterns are present in phonological representations (Dresher 2009: 74).

I assume that all phonological categories are privative: that only one set of two contrastive categories is marked with a label, or feature (e.g. Avery and Idsardi 2001; Iverson and Salmons 1995). For laryngeal features, a two-way privative contrast specifies either the VOICELESS or the VOICED set. A fortis-lenis language like English specifies the VOICELESS series; a voicing language, like French, specifies the VOICED set (Iverson and Salmons 1995). I represent these differences through phonemic aspiration in English VOICELESS phonemes (e.g. /p^h, t^h, k^h/) and plain in English VOICED phonemes (e.g. /p, t, k/), as in /p^hɪn/ 'pin' and /pɪn/ 'bin'. In voicing languages (e.g. French, Portuguese, Dutch, Tagalog), the oppositions are represented by VOICELESS /p, t, k/ and VOICED /b, d, g/, e.g. Tagalog /káma/ 'bed' and /básа/ 'read'. The sets that do not receive phonological features are 'unspecified'.

Because the phonology operates on symbolic categories (Keating 1996: 263), they are unpronounceable abstractions. The categories are matched (or "completed") with articulatory gestures to become pronounceable (Avery and Idsardi 2001: 46–47). In the sound system presented here, gestures are added at the Phonetic-Phonological level, an intermediate stage between the phonology and the phonetics (Purnell and Raimy 2015: 527). For instance, English VOICELESS obstruents are completed with the gesture [spread] for aspiration. On the other hand, French VOICED obstruents receive the gesture [slack] for voicing. The Phonological level comprises abstract labels, while the Phonetic-Phonological level consists solely of articulatory gestures (Purnell and Raimy 2015: 527).

Just as specified phonological categories are completed with articulatory gestures at the Phonetic-Phonological level, so may gestures be added to unspecified phonological

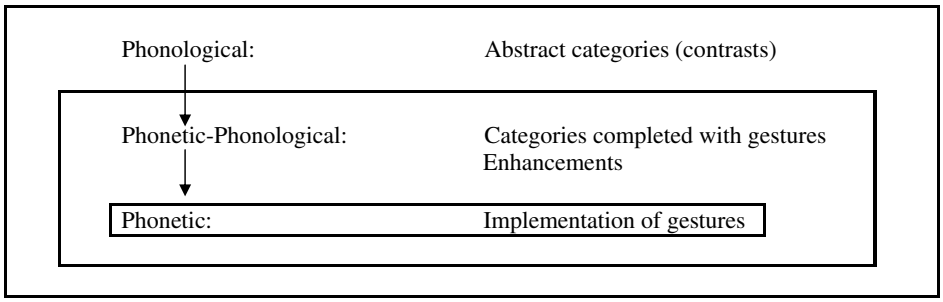


Fig. 8.1: Levels of representation in the sound system

categories as an 'enhancement' (Keyser and Stevens 2006). Enhancing gestures make unspecified categories more easily distinguishable from the specified ones; these gestures also differ from those that correspond to phonological categories by being subject to regional, social, and personal variations (Keyser and Stevens 2006: 38–40).

The Phonetic level of representation governs the implementation of gestures, i.e. their conversion to physical units that "set various aspects of the representations such as amplitude, attack, duration, etc." (Purnell and Raimy 2015: 543). In contrast to the Phonological and Phonetic-phonological levels, Phonetic representations are continuous realizations that exist in time and space (Keating 1996: 263). These include the amount of prevoicing in voiced obstruents, voice onset time (VOT) in aspirated obstruents, and/or sonorant formant values, for example. Figure 8.1 summarizes these levels of representation. Transfer of the Phonological and Phonetic-phonological levels implies the transfer of dependent modules. Transfer of the Phonetic level, on the other hand, need not result in changes to phonological categories or gestures, just in how those categories and gestures present in speech.

Many gradient acoustic and perceptual cues relate to phonological contrasts, so that phonetics and phonology have a non-linear relationship (Purnell 2009). I assume that these asymmetries apply to contact situations as well. Although some argue against a definitive separation between phonetics and phonology (e.g. Ohala 1990), contact data demonstrate the transfer of a section of the sound system (Section 4). Figure 8.1 sketches a testable framework for evaluating the phonological and phonetic effects of languages in contact, yet still allows researchers to draw on phonetic data to interpret and evaluate phonological questions (Ohala 1990: 168).

4. Contact and modularity

In van Coetsem (1988), phonological imposition (SL agentivity) involves the transfer of all three levels of representation from the dominant SL to the non-dominant RL. However, some impositions target a portion of the sound system. I explore two broad types of imposition. First, a systemic imposition, where the entire phonological component – with gestural completions, enhancements, and rules governing the implementation of gestures – is transferred through SL agentivity. Second, a local imposition, where only

a part of the sound system is transferred, I argue, also through SL agentivity due to increased use/proficiency (local dominance) within that module of the sound system. Local impositions may target the Phonetic or Phonetic-Phonological levels, and may also lead to bidirectional influences where each language is both the RL and SL (i.e. convergence; van Coetsem 1988: 87).

I explore the interactions of voicing languages with the fortis-lenis system in English, drawing on examples from French, Portuguese, Dutch, and Tagalog. To compare the abstract categories across systems, VOICELESS refers to specified aspirated /p^h, t^h, k^h/ in English and unspecified voiceless /p, t, k/ in French, Portuguese, Dutch, and Tagalog. VOICED, on the other hand refers to unspecified plain-voiceless /p, t, k/ in English and specified voiced /b, d, g/ in French, Portuguese, Dutch, and Tagalog. Even though the lenis VOICED and voiceless VOICELESS segments (i.e. /p, t, k/) are acoustically similar, the opposition is maintained between languages with different laryngeal features (Uffmann 2015: 255).

4.1. Systemic imposition

A systemic imposition is the type of transfer van Coetsem (1988) defines through SL agentivity where a bilingual uses the phonological system of the dominant language to produce the sounds of the less dominant one. Systemic impositions target the Phonological level of representation, imposing not only abstract categories, but also the subordinate levels − Phonetic-Phonological and Phonetic − that convert those categories into pronounceable forms, realized in the speech signal. Van Coetsem (1988: 11−12) discusses a French-dominant bilingual speaking English without aspiration, as in '[p]it' for '[p^h]it'. The French-speaker learns that English has a VOICED-VOICELESS opposition but transfers the French system for marking those contrasts into English, as in (8.2), with concentric boxes delimiting the three levels of representation in Figure 8.1. Again, French marks VOICED /b, d, g/ because voiceless stops become voiced in French laryngeal assimilations (Iverson and Salmons 2003: 6). The VOICED category in French is phonologically active and, therefore, marked in the French representational system (Dresher 2009).

In Figure 8.2, the French Phonological level of representation is transferred to English through SL agentivity; French is the dominant SL and English the RL (represented by solid and dotted boxes, respectively). The abstract contrast is transferred along with the subsequent levels of representation: the completion of the VOICED series with the gesture [slack] (Phonetic-Phonological level), and the implementation of that gesture (Phonetic level). Systemic impositions are the type of phonological transfer produced at a stage of emerging bilingualism where forms in the RL are systematically adjusted to SL patterns (Matras 2009: 224−226). Systemic impositions include substitutions for RL segments absent from the SL form, as in Japanese /b/ for English /v/ in *yunibashiti* 'university' and *terebijon* 'television' (Matras 2009: 169). Furthermore, Korean transfers a phonological rule whereby /s/ is palatalized to [ʃ] before /i/ in English words like '[ʃ]ick' and '[ʃ]ing' (cf. 'sick', 'sing'), yet surfaces as [s] in other contexts as in '[s]oap' (Idsardi and Son 2004: 53). Systemic impositions, then, consist of the reproduction of

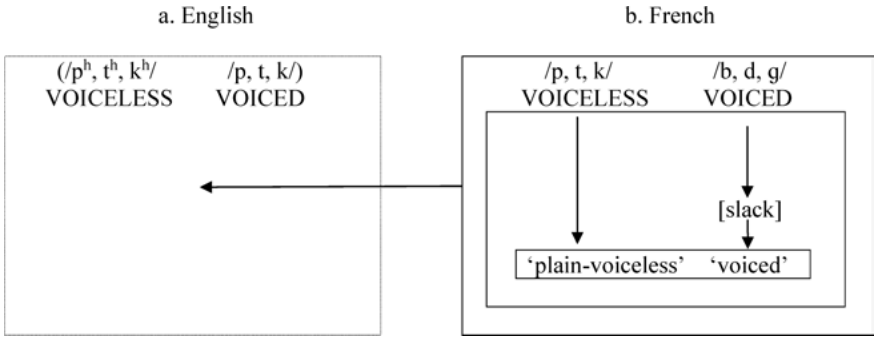


Fig. 8.2: Systemic imposition

the RL using SL phonological categories and phonological rules, and the processes for producing those categories in speech.

4.2. Local impositions

As L2 proficiency increases, dominance relationships may shift and a speaker may "impose material from this language upon his native language" (van Coetsem 1988: 76). English-dominant bilinguals living in Brazil, for example, show Portuguese influence on their English VOTs (i.e. aspiration) in the VOICELESS set with a statistically significant correlation between Portuguese proficiency and low VOT (less aspiration) in casual speaking contexts (Major 1992: 197). Increased use and processing of Portuguese results in more dominance of Portuguese phonetics relative to English, without imposition of the Portuguese VOICING categories, as in Figure 8.3. Here, although English remained the dominant language for many speakers, Portuguese is the source of influence on the phonetic VOT targets in the English RL. Because some English aspiration was preserved, speakers maintain the English contrastive system that marks VOICELESS at the Phonological level and completes that category with [spread] at the Phonetic-Phonological level. Portuguese, however, influences the realization of that gesture in a continuous, gradient manner − a decrease in the length of aspiration in real time. This type of local imposition does not introduce novel features into the RL, but results in adjustments of the physical properties of the gestures that are part of the RL's sound system.

Like French and Portuguese, Dutch has phonologically voiced stops (Iverson and Salmons 2003). In her analysis of adult L1 Dutch speakers' acquisition of the English VOICING contrast in stops, Simon (2011: 239) finds that "the participants had learned to produce aspiration, but had not learned to produce short-lag stops in English". The bilinguals produce English VOICED /p, t/ with prevoicing, indicating the local imposition of the [slack] gesture from the Dutch VOICED /b, d/, as in Figure 8.4. VOICELESS labials and coronals are similar to monolingual targets with respect to VOT in each language, with long-lag in English and short-lag in Dutch. (Dutch lacks a VOICED dorsal stop.)

The native-like productions of English /t^h/ and Dutch /d/ suggest that the participants acquired each language's laryngeal system (English aspiration, Dutch voicing). The im-

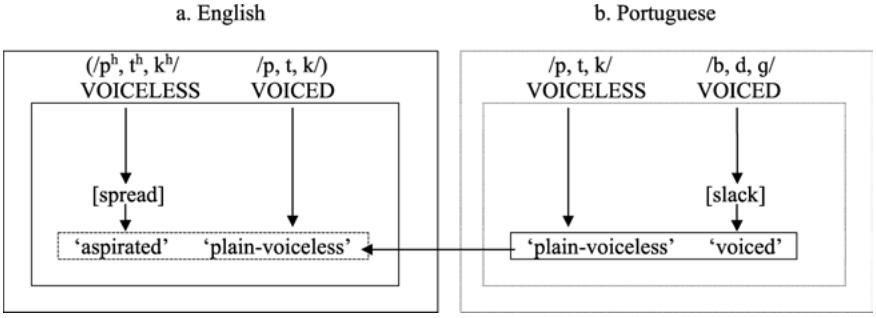


Fig. 8.3: Local imposition (Phonetic level)

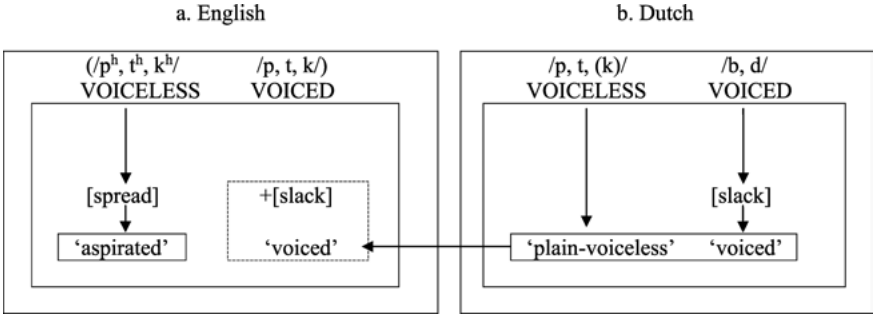


Fig. 8.4: Local imposition (Phonetic-Phonological level)

posed gesture from Dutch targets the phonologically unspecified VOICED /p, t/ in English. The effect enhances the contrast between VOICED and VOICELESS in the English of bilinguals because the perceptual distance between [p^h, t^h] and [b, d] is greater than that between [p^h, t^h] and [p, t]. The transfer thus does not require adjusting the English phonology; it merely adds a gesture to the unspecified (VOICED) member. Because the Phonetic-Phonological level operates on gestures, the inclusion of this new feature indicates a transfer from this level of representation. Furthermore, the direction of transfer from Dutch to English indicates Dutch SL agentivity and dominance, though not to the extent that it dissolves the English categories.

While the Dutch-English bilinguals discussed above maintain separate phonetic systems, Kang, George, and Soo (2016) show bidirectional phonetic influence between English-Tagalog bilinguals. They compare words such as English aspirated 'carpet' and voiceless 'goblin' to Tagalog voiceless *káma* 'bed' and voiced *básа* 'read' (Kang, George, and Soo 2016: 195–197). Their data show marginally statistically significant VOT increases for VOICELESS stops in Tagalog, as well as significant differences in voicing for VOICED stops in both languages. The Tagalog bilinguals likely acquired both types of contrasts since each language "faithfully mirrored native comparisons" (Kang, George, and Soo 2016: 202). The bidirectionality of influences from English and Tagalog, especially in the VOICED consonants, suggests an even linguistic dominance relationship. In such cases, RL agentivity and SL agentivity co-occur in relation to the same language and "*it is precisely this interaction with the lack of a clear distinction between the two*

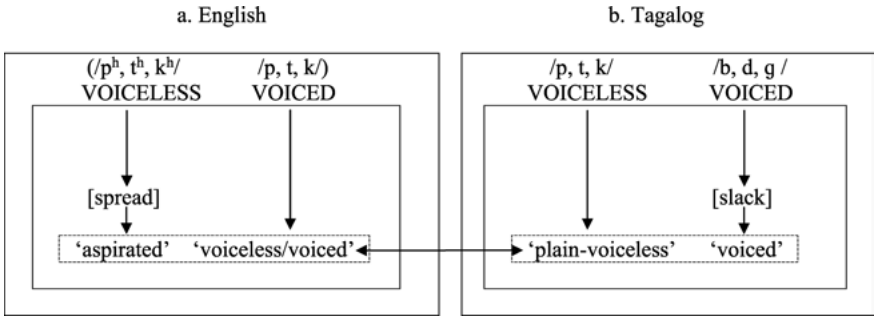


Fig. 8.5: Bidirectional local imposition (convergence)

transfer types that strongly favors a form of convergence of the contacting languages" (van Coetsem 1988: 87, original emphasis). This convergence is represented in Figure 8.5, where English and Tagalog are both SL and RL. However, their phonetics influence each other asymmetrically. Evidence shows "an equivalence classification of English and Tagalog voiced stops in bilinguals' speech, but the retention of separate *phonetic* [emphasis mine] categories in their voiceless stops" (Kang, George, and Soo 2016: 209). Influence on the phonetic − and not the phonological − categories supports the local nature of this bidirectional imposition.

English and Tagalog show statistically significant differences in the VOICED consonants, approaching each other's voicing targets: English stops were more prevoiced and Tagalog stops were less prevoiced than control groups (Kang, George, and Soo 2016: 207). As in the case of Dutch-English bilinguals (cf. Figure 8.4), prevoicing (the phonetic reflex of the VOICED category in Tagalog) enhances the English lenis-fortis contrast. Each language's phonological contrasts are preserved with gradient changes to the phonetic implementation of gestures, and possibly with further changes to the English Phonetic-Phonological level through the introduction of a VOICED enhancement. While the English-Tagalog bilinguals have greater absolute VOTs for VOICELESS stops (Kang, George, and Soo 2016: 202), these differences may reflect the results of a local imposition from English, without causing the Tagalog of bilinguals to differ systemically from that of monolinguals. This model predicts such influence from English as linguistic dominance in Tagalog decreases relative to English.

The asymmetric VOICING patterns of English-Tagalog bilinguals reveal the need to investigate how features introduced through contact interact with the broader phonological system, and how they are adopted and spread throughout a speech community. More detailed analyses of the relationship between linguistic dominance and its effects on local impositions are also warranted; that bidirectional influences from uncertain, or relatively even, dominance relationships obscure and merge the two transfer types identifies an avenue for further inquiry.

4.3. Implications

Both systemic and local impositions introduce new forms to RL sound systems, and each type of imposition in Figures 8.2−8.5 occurs at different stages of individual bilin-

gualism. These processes support a direct relationship between intensity of bilingualism and overlap of phonological systems (Matras 2009: 226). Systemic impositions represent the lowest degree of bilingualism and the most separation between systems; bidirectional local impositions reflect the highest degree of bilingualism and (partial) overlap of the phonetic systems, yet with distinction between phonological systems. Changes in the phonetic productions of phonological categories brought on by local impositions produce different results at different stages of L2 acquisition and proficiency. The extent to which speakers maintain separate systems is likely the result of continued use of each language. Furthermore, it appears that L2 acquisition develops 'bottom up', beginning with the Phonetic level, then proceeding to the Phonetic-Phonological level, and finally to the Phonological level. The imposition of a level of representation, then, carries with it any subordinate level. The case studies further suggest that more consistent use is required to keep the phonetic levels distinct than is necessary for the phonological levels. More research is required for both these claims, and there is likely great individual variation, particularly with the latter. However, this model accounts for how variability in the use of each language across the lifespan affects the degree to which bilinguals can maintain separate sound systems and levels of representation. The analysis informs research on interlanguage grammar, and its role in restructuring phonological systems (Eckman and Iverson 2015), and benefits work on contact-induced change by detailing the mechanisms by which different types of features (e.g. phonological categories, articulatory gestures, or phonetic targets) from one language enter the system of another.

5. Conclusion

I have presented a set of interconnected models for investigating processes and effects of phonological and phonetic contact. Portions of the sound system are targets for imposition, dependent upon relative linguistic dominance, by refining descriptions of imposition: systemic impositions target the phonology; local impositions target sub-phonological levels of representation, and may involve bidirectional transfer without shifts in overall linguistic dominance. I show how transfers of different types of phonetic and phonological material occur in contact situations. Transfers targeting the Phonological level involve the imposition of one language's phonological categories and rules on the other. Transfers at the Phonetic-Phonological level introduce new gestures into the RL, but do not alter the RL's phonological representations. Finally, Phonetic level transfers result in physical modifications of RL gestures.

The integration of phonological theory with contact theory contributes to our understanding of a range of phonological and phonetic contact effects by specifying directions for improving that understanding with detailed analyses of proficiency and dominance relationships, and how they relate to fine-grained phonetic influences across languages. This model outlines multiple types of influences on sound systems at three different levels of representation. While this analysis has focused on the synchronic processes and outcomes of sound systems in contact, the model further lends itself to examinations of contact-induced change.

Acknowledgments

I would like to thank Matthew Greene, Mirva Johnson, Samantha Litty, Joe Salmons, and two anonymous reviewers for their insights and advice on previous versions of this chapter. All errors and oversights are my own.

6. References

Avery, Peter & William Idsardi
2001 Laryngeal dimensions, completion and enhancement. In T. Alan Hall (ed.), *Distinctive feature theory*, 41–70. Berlin: Mouton de Gruyter.

Benmamoun, Elabbas, Silvina Montrul & Maria Polinsky
2013 Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics* 39. 129–181.

Docherty, Gerard J.
1992 *The timing of voicing in British English obstruents*. Berlin: Foris.

Dresher, B. Elan
2009 *The contrastive hierarchy in phonology.* Cambridge: Cambridge University Press.

Dubois, Sylvie & Barbara Horvath
1998 Let's tink about dat: Interdental fricatives in Cajun English. *Language Variation and Change* 10. 245–261.

Eckman, Fred & Gregory Iverson
2015 Second language acquisition and phonological change. In Patrick Honeybone and Joseph Salmons (eds.), *The Oxford handbook of historical phonology*, 637–643. Oxford: Oxford University Press.

Flege, James E.
1995 Second language speech learning: Theory, findings, and problems. In Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 233–277. Timonium, MD: York Press.

Goldsmith, John A.
1995 Phonological theory. In John A. Goldsmith (ed.), *The Handbook of phonological theory*, 1–23. Cambridge: Blackwell.

Haugen, Einar
1950 The analysis of linguistic borrowing. *Language* 26(2). 210–231.

Howell, Robert B.
1993 German immigration and the development of regional variants of American English: Using contact theory to discover our roots. In Joseph Salmons (ed.), *The German language in America*, 188–217. Madison: Max Kade Institute.

Idsardi, William & Sun Ah Son
2004 What do Korean speakers learn about English fricatives? In Susumu Kono, Ik-Hwan Lee, John Whitman, Joan Maling, Young-Se Kan & Young-joo Kim (eds.), *Harvard studies in Korean linguistics* 10, 48–59. Cambridge, MA: Deparment of Linguistics, Harvard University.

Iverson, Gregory & Joseph Salmons
1995 Aspiration and laryngeal representation in Germanic. *Phonology* 12(3). 369–396.

Iverson, Gregory & Joseph Salmons
2003 Legacy specification in the laryngeal phonology of Dutch. *Journal of Germanic Linguistics* 15(1). 1–26.

Kang, Yoonjung, Sneha George & Rachel Soo
2016 Cross-language influence in the stop voicing contrast in heritage Tagalog. *Heritage Language Journal* 13(2). 184–218.
Keating, Patricia A.
1996 The phonology-phonetics interface. In Ursula Kleinhenz (ed.), *Interfaces in phonology*, 262–278. Berlin: Akademie Verlag.
Keyser, Samuel Jay & Kenneth Noble Stevens
2006 Enhancement and overlap in the speech chain. *Language* 82(1). 33–63.
Major, Roy C.
1992 Losing English as a first language. *The Modern Language Journal* 76. 190–208.
Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.
Natvig, David
2017 A model of underspecified recognition for phonological integrations: English loanwords in American Norwegian. *Journal of Language Contact* 10(1). 22–55.
Ohala, John J.
1990 There is no interface between phonology and phonetics: A personal view. *Journal of Phonetics* 18(2). 153–171.
Purnell, Thomas
2009 Phonetic influence on phonological operations. In Eric Raimy & Charles Cairns (eds.), *Contemporary views on architecture and representations in phonology*, 337–354. Cambridge: MIT Press.
Purnell, Thomas & Eric Raimy
2015 Distinctive features, levels of representation and historical phonology. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford handbook of historical phonology*, 522–544. Oxford: Oxford University Press.
Purnell, Thomas, Joseph Salmons, Dilara Tepeli & Jennifer Mercer
2005 Structured hetereogeneity in laryngeal phonetics: Upper Midwestern final obstruents. *Journal of English Linguistics* 33(4). 307–338.
Putnam, Michael T., Silvia Perez-Cortes & Liliana Sánchez
To appear Language attrition and the feature reassembly hypothesis. In Barbara Köpke & Monika Schmid (eds.), *The Oxford handbook of language attrition.* Oxford: Oxford University Press.
Sharma, Devyani
2011 Style repertoire and social change in British Asian English. *Journal of Sociolinguistics* 15(4). 464–492.
Simon, Ellen
2011 Laryngeal stop systems in contact: Connecting present-day acquisition findings and historical contact hypotheses. *Diachronica* 28(2). 225–254.
Uffmann, Christian
2015 Loanword adaption. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford handbook of historical phonology*, 644–665. Oxford: Oxford University Press.
van Coetsem, Frans
1988 *Loan phonology and the two transfer types in language contact.* Providence, RI: Foris.
Weinreich, Uriel
1966 [1953] *Languages in contact: Findings and problems*, 4th edn. The Hague: Mouton.
Winford, Donald
2005 Contact-induced changes: Classification and processes. *Diachronica* 22. 373–427.

David Natvig, Northfield, MN (USA)

9. Pidginization and creolization

1. Introduction
2. Pidginization and creolization
3. Types of pidgins and creoles
4. Social contexts
5. Linguistic properties
6. References

1. Introduction

Although research on pidgins and creoles (P/Cs) started before the middle of the 20th century, many view the 1959 and 1961 Mona, Jamaica conferences, and the volume edited by Hymes (1971b), as the main catalysts for bringing together researchers and for bringing P/Cs "to the wider attention of linguists" (Kouwenberg and Singler 2008: 3). Since its inception, the field has been dominated by debates about the origin and development of P/Cs. This research, and the rise of contact linguistics as a separate field of investigation since the 1980s, has been instrumental in vindicating P/Cs, and the processes of their emergence, of the 'shamefulness' pinned on them by Eurocentric research paradigms.

P/Cs constitute a continuum, in that the known cases are characterized by varying degrees of similarity and difference. One reason is the terms' lay origin. *Pidgin* probably comes from English 'business', being first applied to Chinese Pidgin English (Baker and Mühlhäusler 1990). 'Creole' derives from Spanish *criollo*, and initially referred to Spanish-speaking people of European origin born in the Caribbean and/or South America before being extended to designate new language varieties that emerged in such contexts. Another reason is the definitive criteria for each category. Some researchers focus on socio-historical and social criteria (Mufwene 2000), while others argue that they are linguistically distinct (McWhorter 2018). There is growing consensus now that neither suffice to fully define them.

Historical linguistics research has traditionally treated P/Cs, and their processes of emergence, as exceptional or marginal, calling them "abnormal" language transmission (Thomason and Kaufman 1988: 146), because their grammars were not inherited from a single source language. Challenging this view, Winford (2003) and Mufwene (2001) point to the close similarities between the processes and outcomes of change involved in the emergence of P/Cs and in other types of contact. DeGraff (2003) also shows that their features are often more complex than those found in their input languages.

2. Pidginization and creolization

Early theories and e.g. McWhorter (2018) see P/Cs as intricately linked. Hall's (1962) life cycle theory argues that creoles developed from pidgins, which in turn originate in jargons. People who did not share a common language deployed idiosyncratic strategies of communication to suit their immediate communication needs. When they stabilized and conventionalized, a pidgin emerged. When a pidgin became the main means of communication of a group and expanded its expressive options, a creole was born. Al-

https://doi.org/10.1515/9783110435351-009

though this developmental model makes untenable assumptions of social and linguistic homogeneity across contexts and categories, it nevertheless highlights important aspects of these processes: pidginization is conceptualized as a process of linguistic reduction leading to a reduced means of communication, while creolization entails expansion giving rise to a full-fledged language. Both processes consist of several subprocesses (Hymes 1971a), and this multi-processual nature is highlighted with terms such as P/C formation, or genesis.

There are some commonalities between them. Their formation involves contact between speakers with different linguistic repertoires. In both cases, retention of material from the creators' first languages (L1s), substrate influences, and innovations (the creation of new structural patterns not shared by the input languages) play a role. However, pidgins are linked to processes of simplification and innovation, while creole genesis research has highlighted the contribution of the input languages.

2.1. Pidgin formation

Pidgin genesis involves three linguistic processes: simplification, retention, and innovation (Bakker 2008; Bickerton 1981; Winford 2003). Discussions have particularly focused on simplification (Trudgill 2009), pointing to phenomena such as small vocabularies and phonological inventories, lack of bound morphology, and a small number of structural rules. The Russian-Norwegian Pidgin Russenorsk, for instance, had a vocabulary of between 150–200 words, consisting mostly of content words (nouns, verbs, adjectives) and two main pronouns (Winford 2003: 273). Other pidgins have larger lexica, but only a small number of functional categories, such as prepositions, cross-clausal markers, determiners, and temporal markers. Lexical items are often polysemous and categorically under-specified (Drechsel 1996). The vowel and consonant inventories are smaller than those of the source languages, and there are fewer phonological contrasts. The most debated feature is bound morphology. Pidgins have few (or no) derivational and inflectional affixes. Those that exist often developed through processes of grammaticalization. Crowley (2008: 77) also notes the absence of redundancy, morphological irregularity, and analytical structures, and a small number of structural patterns at the levels of syntax, morphology, word order, and low derivational depth. Reduced structures arise due to people's attempts to make sense of each other's languages, involving language learning and reduction of people's primary languages.

Pidgin grammars also retain features from their input languages which emerged through substrate influence or transfer. This refers to the use of material from the linguistic repertoire of their main creators, such as Africans (Atlantic creoles) and Pacific Islanders (Pacific creoles). Aside from lexical items and phonological segments, it involves structural patterns and principles from all levels of grammar, including specific meanings and functions. Superstrate influence refers to the adoption of forms from the sociopolitically dominant language. Both influences play a restricted role in pidgin genesis. Although pidgins retain lexical items from both sources (Bakker 2008: 137), the realization of vowels and consonants, and certain phonotactic constraints, come from their creators' first languages (L1s) (Baker and Mühlhäusler 1990), and may also be cross-linguistically marked. Retention of structural patterns varies considerably. In the case of

the Ndyuka-Trio Pidgin (Huttar and Velantie 1997), most of its syntax comes from Trio, while in other cases, such as Delaware Pidgin, L1 input seems to have contributed less.

Innovations are features that are not directly derived from a source language. For instance, a language might extend its expressive capacity, or range, by developing new morphological patterns, such as iteration, to express iterative aspects or large quantities, or new morphological markers, due to processes of grammaticalization, pragmatic reinterpretation, and semantic extension or reduction. Delaware Pidgin uses innovative compounding patterns to create new lexical items (Goddard 1997), Russenorsk developed a verbal marker, and Sankoff and Brown (1976) demonstrate how syntactic constructions emerged in Tok Pisin from discursive patterns. Some languages also adopt new word orders, or develop new constituent orders.

It is also widely accepted that pidgin formation is regulated by similar cognitive and operating principles as (early) second language acquisition (SLA), such as a focus on semantic transparency (Winford 2003: 225–227).

2.2. Creole formation

Creoles are full-fledged languages that serve all communicative needs of a community. It is generally agreed that their emergence involves substrate, superstrate, and adstrate influences and processes of innovation. Debates have focused on the relative contributions of each source, the nature of the linguistic (sub)processes, the agents of change, and the duration of the process. Bickerton's (1981) bioprogram hypothesis argued that creole formation is a type of L1 acquisition, in which children rely on a blueprint for language. This view has been successfully challenged by subsequent research. Research on social history (Arends 1995; Roberts 2000; Singler 1996) challenged this social scenario by pointing to the lack of children, the agency of adults, and the relative homogeneity of the adults' L1s as important factors that determined the impact of substrate influence. In other cases, such as Hawaiian Creole, the substrate input was quite heterogeneous (Siegel 2008), but substrate influence still occurred. Jourdan and Keesing (1997) and Arends (1993) also highlight, albeit from different perspectives, that creole grammars emerged gradually and not always linearly, over a longer period of time.

Much of the research focused on the role of their creators' L1s. According to Lefebvre's (1998) relexification account, creators of a creole retained the structural information from their native lexical items and linked them to semantically overlapping forms in the superstrate language, thereby creating new lexical items that have superstrate phonological shapes and semantic, syntactic and distributional features mostly derived from the substrates. Others (Siegel 2008; Winford 2003) maintain that creole formation is similar to transfer in early SLA: the creators identified (seemingly) semantically or functionally similar lexical items in varieties of the dominant language and their L1s, and projected the properties of the latter onto the former. This gave rise to mixed lexical items and constructions that have properties shared by all input varieties, and those that are unique to one or the other. Many instances have been documented where the lexical shapes of functional items and some of their properties, derive from the superstrate, while the bulk of the semantic and syntactic information overlaps either only with the substrate or in part with both the substrate and superstrate counterpart. A case in point are the copula elements (*n*)*a* and *de* in the Eastern Maroon Creoles (EMC). They derive from the distal

demonstrative *that* and the locative adverb *there*, respectively, and their distributions are very similar to their Gbe counterparts (Migge 2002). This probably occurred during creole formation, and adults were the main agents. Siegel (2008), however, proposes, on the basis of research on Pacific varieties, that L1 transfer occurred when bilingual children adopted the emerging variety as their main in-group language. Studies provide evidence of the retention of abstract structural patterns and principles, of semantic and syntactic information, and of lexical items (e.g. papers in Essegbey, Migge, and Winford 2013; Keesing 1988; Lefebvre 1998; Migge and Smith 2007; Muysken and Smith 2016; Siegel 2008). More work is required on the role of discourse (Meyerhoff 2008) and pragmatic phenomena. Substrate influence clearly played an important role in the emergence of many creoles. However, it did not affect them to the same degree, or in the same way due to differences in typology and in the contact setting.

Superstrate influence has received less attention. It generally contributed the bulk of the lexicon, which was then modified to varying degrees under L1 influence and universal processes. Superstrates also provided word order phenomena, as in SLA (Plag 2008) and certain kinds of constructions, such as aspectual or demonstrative modifier constructions, that are still found in regional and social varieties (Lalla and D'Costa 1990). The creators acquired (aspects of reduced) varieties of the dominant language and restructured them, because they lacked sufficient access and/or motivation to acquire them fully (Baker 1990). Chaudenson (2001) posits that the first generation of agents in creole genesis learned relatively rich L2 varieties of the superstrate language. Over time, the varieties became targets of SLA for new arrivals and underwent change. The role of the superstrate was not homogeneous across contexts at emergence or following it, and thus different languages display different degrees of superstrate influence.

While few researchers nowadays support the bioprogram hypothesis (Veenstra 2008), is it accepted that universals of language contributed to creole formation. Winford (2003: 347) suggests that there are universals that act as constraints on substrate and superstrate influence, and on internally motivated changes. The former are cognitive processing constraints, and are usually couched in terms of principles, such as: L1 retention is facilitated if constructions are structurally congruent with those in the superstrate. (1) illustrates an EMC existential construction which is derived from a locational construction.

(1) *wagi de.* < *car Ø there*
car there
'There is/exists a car.'

These processes operate concurrently with other processes. In contrast, internally-motivated processes are pragmatically driven, and operate after substrate and superstrate influence have occurred, affecting the elements and constructions they gave rise to. There is evidence that the past time marker *be*(*n*) in the EMC emerged due to grammaticalization (Migge and Goury 2008), while the habitual function of the imperfective marker *de* and the nominal copula function of the existential verb *de* in the Surinamese Creoles appear to have developed from functional extension (Migge 2002).

3. Types of pidgins and creoles

Traditionally, P/Cs have been classified according to their main lexical input language(s), designated as English(-based), Portuguese(-based), etc. P/Cs. Increasingly, researchers focus on socio-historical and usage-based criteria. On the one hand, there are prototypical pidgins. Following Hymes (1971a) and Holm (1988), Winford (2003: 270) defines them as varieties of language that are "characterized by highly reduced vocabulary and structure, which are native to no one, and serve as lingua francas for certain restricted communicative functions such as trade". Prototypical cases are Delaware Pidgin, Russenorsk, and Chinook Jargon/Pidgin.

Languages that are locally referred to as pidgins, such as Kamtok (Cameroon) or Solomon Islands Pijin, but defy these criteria, are referred to as elaborated or expanded, pidgins (Winford 2003), or pidgincreoles (Bakker 2008). They are languages that started out as a restricted-purpose means of communication and subsequently significantly expanded their domains of usage, grammar, and communicative strategies. A well-documented case is Melanesian Pidgin, and its daughter languages, Tok Pisin, Solomons Pijin, and Bislama. They still coexist with their substrate languages, and have various functions − as an interethnic means of communication, a means of status negotiation, and a main means of communication. Although they are not L1s, they have developed into full-fledged languages, fulfilling all of their speakers' communicative needs. Bakker (2008) includes Pacific and African varieties, such as the Melanesian varieties and Nigerian Pidgin English, along with Sango, Kituba, Lingala and Arabic varieties (Tosco and Manfredi 2013), in this category. Winford (2003: 288) distinguishes the latter from the former, in that "they are really somewhat simplified versions of one source language […]".

Creoles are inextricably tied up with the plantation context. There are two types of creoles: radial and intermediate creoles (Winford 2003: 208). Radical creoles emerged in contexts with comparatively little intense contact with the socially dominant European language, such as the Suriname Creoles. They diverge structurally in significant ways from the European input languages. Their lexicon derives from the European languages, but the meanings and distributions of lexical items only overlap in part with their English counterparts. In contrast, intermediate creoles, such as Bajan, develop under sustained influence from the socially dominant language, and structurally and lexically closely resemble its varieties, such as varieties of English. Many of the creoles in the Caribbean region, including South and Middle American Creoles, probably fall somewhere between these two broad categories of creoles.

4. Social contexts

An important reason for the differences between (and among) P/Cs are the differences in the contexts of their emergence. Prototypical pidgins emerged in interethnic contact settings which facilitate interaction (Versteegh 2008). Trading activities have provided an important context, but other settings, such as multilingual crews, military invasions (Reinecke 1971), and domestic work contexts (Hosali 1992) are also attested. More recent work focuses on Arabic Pidgins that have arisen in interactions both between

citizens of the Gulf States and expatriate Asian and African workers, and among the latter group (Almoaily 2014). Many of the expatriates work in the domestic sphere (Bizri 2014), which involves exposure to foreigner talk or L2 varieties of Arabic with little sustained contact with L1 Arabic. The Maritime Pidgin referred to as Pacific Pidgin (English) (Clark 1979), which arose among sailors working on ships and crews in the late 18th and early 19th century in Pacific harbors, is a good example of the crew context. It facilitated work-related, but not strongly hierarchical, interactions between people with different linguistic backgrounds. The trade context includes Trio-Ndyuka Pidgin (Huttar and Velantie 1997), which arose in the interior of Suriname to facilitate trading interactions between Maroons and Amerindians. Classic cases that have been widely discussed are Russenorsk, which arose at the end of the 19th century among Russian fishermen and Norwegian merchants in Norway, and various North American Pidgins, such as Chinook Jargon, Mobilian Jargon, and Delaware Pidgin (Drechsel 1996; Goddard 1997). These initially served inter-Amerindian trade contacts, and later Amerindian-European commercial interactions.

Expanded pidgins often also originated in the above contexts, and were then extended to new contexts and expanded linguistically. The Pacific Pidgin that had emerged during the Sandalwood trade and maritime activities in the Pacific in the 18th and 19th century was an important input for Melanesian Pidgin English, which became an important means of interethnic communication among Pacific islanders on the 19th century Queensland and Samoan plantations. When workers returned home, it was spread to the different islands, becoming an interethnic means of communication. With increasing urbanization, it is now acquiring first language users and expanding into the private sphere. Comparable processes have taken place in other contexts, such as western Africa.

Radical and intermediate creoles both emerged in the plantation context of the Caribbean Sea and Indian Ocean. Prototypical pidgins probably served as inputs, but these creoles are most likely not elaborations of the former, as claimed by supporters of the Afro-genesis view of creole formation (McWhorter 1995). Instead, enslaved Africans were exposed to a range of varieties of the dominant European languages, which they then reinterpreted based on their own linguistic background to construct a common means of communication. Initially, creoles existed side-by-side with African languages, but quickly became quite dominant. Their use as mother tongues probably only started when the population began to sustain itself, and a local identity emerged. Radical creoles emerged on the worker-intensive sugar plantations, where access to varieties of the dominant language was limited. The ratio of Europeans to slaves was often very low (1:10+) because the workforce consisted mostly of enslaved Africans, and interactions with Europeans were mostly superficial and infrequent. Radical creoles formed primarily in interactions among forced laborers, such as among Africans, and between Africans and European indentured laborers, rather than in master-enslaved encounters. The forced labor context also did not provide incentives to all to learn the dominant language, or to associate with its speakers (Baker 1990). However, there were important differences between plantation settings (Mintz 1971; Singler 2008). In some settings, the ratios were much higher and/or the crops were less worker-intensive, allowing more intense contact between laborers and colonizers, resulting in increased input from the colonizers' languages due to more frequent interactions.

5. Linguistic properties

P/Cs are socio-historically and socially heterogeneous, and do not make up a linguistic class, but see Bakker et al. (2011). The most explicit claims about structural characteristics have so far been made by Bickerton (1981). According to him, prototypical creoles can be identified on the basis of twelve structural features. Although not tenable, research on creoles has since spilled a lot of ink researching these features, to the detriment of others. More recently, McWhorter (1998, 2018), focusing on only three structural features, proposed that creoles are identifiable as a class because they lack inflectional and (rich) derivational morphology, and make little or no use of tonal distinctions to mark grammatical relations. Finally, Bakker (2008; following Seuren 1998) argued that pidgins differ from creoles and expanded pidgins, in that the latter have SVO word order, and tense, aspect, and mood markers (TMA), but generally no bound morphology.

The broad area of morphosyntax has received the greatest amount of attention (see also Chapter 6), and within that, tense and aspect categories. Creoles generally express these categories using pre-verbal markers that often derive from verbal forms. Future time reference is expressed by a form derived from a verb meaning 'to go' (*g*)*o* in Atlantic English-lexified creoles, such as Jamaican Creole, or English 'want' *waan* in Krio and Belize Creole. In French-based varieties, such as Haitian Creole, the future can be expressed by *va* from French *aller*. Past time is expressed by elements derived from past-marked forms of 'to be', such as English *been* as in Sranantongo (*ben*), and in French-based varieties, such as Haitian Creole, for instance, where it is expressed by *t(e)* < French *été* or by the unmarked verb. The present time is realized using a progressive construction involving the preverbal imperfective marker (*d*)e < English *there* in the Surinamese Creoles, and *ap* from French *après* in French-based Creoles, such as Haitian Creole. While there are broad similarities in function between such items across creoles, they often differ substantially in terms of their precise meanings, uses, lexical categories, and combinatory possibilities. Prototypical pidgins often lack clearly defined categories, and express these meanings using adverbial forms.

Creoles, like many of their substrates, and pidgins generally make little or no use of number, gender, and/or case markings on nouns and verbs, and different lexical categories are generally not distinguished by separate morphological markers. While plural and singular nouns can be distinguished using number words or definite articles, *a osu* versus *den osu* 'the(sg) / the(pl) house' in the Surinamese Creoles, the noun itself is not marked for plurality. Where discursively necessary, sex might also be indicated using nominal juxtaposition, *uman data* 'female medical doctor' (Migge 2001), but it is not redundantly marked on nouns and pronouns. The same applies to pidgins, which often do not have conventionalized options for marking number and sex.

Due to the focus on English-based P/Cs, it was commonly assumed that P/Cs do not have inflectional morphology, but Bakker (2003) shows that such elements exist and are often inherited (or emerge independently) from internal or contact-induced change. Fanakalo has inherited the future prefix *zo-* and the past suffix *-ile* from Zulu. Broome Lugger Pidgin (Malaysia) inherited the question and conditional suffix *-ka* from Japanese, where it functions as an interrogative and indefinite marker. The word formation process that has come to be closely associated with P/Cs is reduplication, however, Bakker (2008) argues that it is mostly absent in pidgins. Reduplication performs a number of different functions in creoles. In the EMC, reduplication of property items can

express that something has not yet reached the target, *lebilebi* 'redish', or if it is selected by the copula *de*, it refers to a (temporary) state (2).

(2) *a* *buuku* *de* *natinati.*
DET(SG) trousers COP wet-wet
'The trousers are in a wet state.'

Reduplication does not have the same functions and productivity in all creoles and its use in these languages correlates to a certain extent with the types found in their inputs (Kouwenberg 2003). Compounding is also a very productive process in creoles (Braun 2009) and pidgins, but remains largely understudied.

Creoles also differ from prototypical pidgins in clausal and cross-clausal syntax. The former have highly conventionalized systems, while the latter do not. One prominent area of research are serial verb constructions (SVCs), in which two (or more) verbal expressions are juxtaposed without connecting markers, and where one of them, usually the second verb, performs grammatical rather than verbal functions. In directional SVCs, the verbs 'go' and 'come' combine with a main verb not marked for directionality to express either movement away from or towards a location (3).

(3) *Tyay* *den* *koosi* *go* *a* *osu.*
carry DET(PL) clothes go LOC house
'Bring the clothes to the house.' (EMC)

Other SVCs, such as that involving the V2 *moo* 'more', are used to express comparison, and V2 *gi* encodes various dative-type relationships (4) (Migge 2003).

(4) *tya* *a* *moni* *kon* *gi* *mi*
bring DET(SG) money come give me
'bring me the money'

The range and types of SVCs correlate with those in the input languages. There is comparatively little research on complementation and particularly coordination. Another area of research is the copular domain. Pidgins often do not overtly express them, and creoles often have quite complex systems. The EMC distinguishes nominal (*n*)*a*) and locational/existential predication *(de*), and property items are verbs. The debate has mostly focused on the emergence of copula constructions in creoles (Baptista 1999).

The areas that have received comparatively little attention are phonology, including supra-segmentals, and vocabulary. Research (Smith 2008) has shown that, especially creoles, often have fewer contrasts than their input languages, and tend to preserve those segments that are common to all inputs. However, P/Cs also preserve typologically marked elements, such as double articulated plosives /gb, kp/. Phonotactic constraints, phonological processes, and supra-segmentals often derive from the input languages as well, and the little research that has occurred has documented influence from the substrate languages. Pidgins differ from creoles, in that their systems are more heavily influenced by the repertoires of their speakers and are often less conventionalized than in the case of creoles. Finally, there has been some debate about the nature of P/C vocabularies. They were traditionally assumed to be small and to come from the lan-

guage of the socially dominant group for pidgins. However, vocabulary often comes from both contributing languages (in the case of two-language contact settings), or predominantly from one or the other language group involved. For creoles, the politically dominant language generally provides the bulk of the lexicon, but other languages, such as those of their main creators and other important groups (adstrate influence), such as Amerindians in the case of New World creoles, often contribute items for specialized domains, such as medicine, flora and fauna, and cultural practices. Influence from the latter languages is also responsible for some of the semantic changes that affected the vocabulary derived from the dominant language. The domains of semantics (Huttar, Essegby, and Ameka 2007; Jourdan 2000) and pragmatics (Mühleisen and Migge 2005) require further attention.

While we have gained a fair amount of knowledge about the processes involved in the emergence of and the nature of P/Cs, this research often focuses on a subset of P/Cs: those that emerged due to European expansion. As new cases in new contexts are explored, we will likely need to, at least in part, revise current theories and assumptions. The *Atlas of Pidgin and Creoles languages* (Michaelis et al. 2013) is one of the tools which has already proven to be very useful in this regard.

6. References

Almoaily, Mohammad
2014 Language variation in Gulf Pidgin Arabic. In Isabelle Buchstaller, Anders Holmberg & Mohammad Almoaily (eds.), *Pidgins and creoles beyond the Africa-Europe encounter*, 57–84. Amsterdam: John Bejamins.

Arends, Jacques
1993 Towards a gradualist model of creolization. In Francis Byrne & John Holm (eds.), *Atlantic meets Pacific: A global view of pidginization and creolization*, 371–380. Amsterdam: John Benjamins.

Arends, Jacques
1995 Demographic factors in the formation of Sranan. In Jacques Arends (ed.), *The early stages of creolization*, 233–285. Amsterdam: John Benjamins.

Baker, Philip
1990 Off target? *Journal of Pidgin and Creole Languages* 5(1). 107–119.

Baker, Philip & Peter Mühlhäusler
1990 From business to pidgin. *Journal of Asian Pacific Communication* 1(1). 87–115.

Bakker, Peter
2003 Pidgin inflectional morphology and its implications for creole morphology. In Ingo Plag (ed.), *Yearbook of morphology: Special section on pidgins and creoles*, 3–33. Dordrecht: Kluwer.

Bakker, Peter
2008 Pidgins versus creoles and pidgincreoles. In Silvia Kouwenberg & John V. Singler (eds.), *The handbook of pidgin and creole studies*, 130–157. Malden: Wiley-Blackwell.

Bakker, Peter, Aymeric Daval-Markussen, Ingo Plag & Mikael Parkvall
2011 Creoles are typologically distinct from non-creoles. *Journal of Pidgin and Creole Languages* 26(1). 5–42.

Baptista, Marlyse
1999 On the nature of the morpheme *e* in Cape Verdean Creole: To be or not to be. In Klaus Zimmermann (ed.), *Lenguas criollas de base lexical española y portuguesa*, 25–47. Frankfurt: Vervuert; Madrid: Bibliotheca Ibero-Americana 66.

Bickerton, Derek
1981 *Roots of language*. Ann Arbor: Karoma.
Bizri, Fida
2014 Unity and diversity across Asian migrant Arabic pidgins in the Middle East. *Journal of Pidgin and Creole Languages* 29(2). 385–409.
Braun, Maria
2009 *Word formation and creolization: The case of early Sranan*. Berlin: De Gruyter.
Chaudenson, Robert
2001 *Creolization of language and culture* (revised in collaboration with Salikoko S. Mufwene). London: Routledge.
Clark, Ross
1979 In search of Beach-la-Mar: Towards a history of Pacific Pidgin. *Te Reo* 22, 3–64.
Crowley, Terry
2008 Pidgin and creole morphology. In Silvia Kouwenberg & John V. Singler (eds.), *The handbook of pidgin and creole studies*, 74–97. Malden, MA: Wiley-Blackwell.
DeGraff, Michel
2003 Against creole exceptionalism. *Language* 79(2). 391–410.
Drechsel, Emanuel J.
1996 Native American contact languages of the contiguous United States. In Stephen Wurm, Peter Mühlhäusler & Darrell T. Tryon (eds.), *Atlas of intercultural communication in the Pacific, Asia and the Americas*, vol. II.2, 1213–1239. Berlin: Mouton de Gruyter.
Essegbey, James, Bettina Migge & Donald Winford (eds.)
2013 Crosslinguistic influence in language creation: Assessing the role of the Gbe languages in the Surinamese creoles. [Special issue]. *Lingua* 129(7).
Goddard, Ives
1997 Pidgin Delaware. In Sarah G. Thomason (ed.), *Contact languages: A wider perspective*, 43–98. Amsterdam: John Benjamins.
Hall, Robert A., Jr
1962 The 'life-cycle' of pidgin languages. *Lingua* 11(1). 151–156.
Holm, John
1988 *Pidgins and creoles: Theory and structure*, vol. 1. Cambridge: Cambridge University Press.
Hosali, Priya
1992 Syntactic peculiarities of Butler English. *South Asian Language Review* 2(2). 58–74.
Huttar, George & Frank J. Velantie
1997 Ndjuka-Trio Pidgin. In Sarah G. Thomason (ed.), *Contact languages: A wider perspective*, 99–124. Amsterdam: John Benjamins.
Huttar, George L., James Essegby & Felix Ameka
2007 Gbe and other West African sources of Suriname creole semantic structures: Implications for creole genesis. *Journal of Pidgin and Creole Languages* 22(1). 57–72.
Hymes, Dell
1971a Introduction: General conceptions of processes. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 65–90. Cambridge: Cambridge University Press.
Hymes, Dell (ed.)
1971b *Pidginization and creolization of languages*. Cambridge: Cambridge University Press.
Jourdan, Christine
2000 "My nephew is my aunt": Features and transformations of kinship terminology in Solomon Islands Pijin. In Jeff Siegel (ed.), *Processes of language contact*, 99–121. Montreal: Fides.
Jourdan, Christine & Roger Keesing
1997 From fisin to pidgin: creolization in process in the Solomon Islands. *Language in Society* 26(3). 401–420.

Keesing, Roger
1988 *Melanesian Pidgin English and the Oceanic substrate*. Stanford: Stanford University Press.
Kouwenberg, Silvia (ed.)
2003 *Twice as meaningful: Reduplication in pidgins, creoles and other contact languages*. London: Battlebridge.
Kouwenberg, Silvia & John V. Singler
2008 Introduction. In Silvia Kouwenberg & John V. Singler (eds.), *The handbook of pidgin and creole studies*, 1–16. Malden: Wiley-Blackwell.
Lalla, Barbara & Jean D'Costa
1990 *Language in exile: Three hundred years of Jamaican Creole*. Tuscaloosa: University of Alabama Press.
Lefebvre, Claire
1998 *Creole genesis and the acquisition of grammar: The case of Haitian Creole*. Cambridge: Cambridge University Press.
McWhorter, John
1995 Sisters under the skin: A case for genetic relationship between the Atlantic English-based Creoles. *Journal of Pidgin and Creole Languages* 10(2). 289–333.
McWhorter, John
1998 Identifying the creole prototype: Vindicating a typological class. *Language* 74(4). 788–818.
McWhorter, John
2018 *The creole debate*. Cambridge: Cambridge University Press.
Meyerhoff, Miriam
2008 Forging Pacific pidgin and creole syntax: Substrate, discourse, and inherent variability. In Silvia Kouwenberg & John V. Singler (eds.), *The handbook of pidgin and creole studies*, 48–73. Malden: Wiley-Blackwell.
Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.)
2013 *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://apics-online.info/ (accessed 12 August 2018).
Migge, Bettina
2001 Communicating gender in the Eastern Maroon Creole. In Marlis Hellinger & Hadumod Bussmann (eds.), *Gender across languages*, 85–104. Amsterdam: John Benjamins.
Migge, Bettina
2002 The origin of the copulas (*d/n*)*a* and *de* in the Eastern Maroon Creole. *Diachronica* 19(1). 81–133.
Migge, Bettina
2003 *Creole formation as language contact: The case of the Suriname Creoles*. Amsterdam: John Benjamins.
Migge, Bettina & Laurence Goury
2008 Between contact and internal development: Towards a multi-layered explanation for the development of the TMA system in the creoles of Suriname. In Susanne Michaelis (eds.), *Roots of creole structure*, 301–331. Amsterdam: John Benjamins.
Migge, Bettina & Norval Smith (eds.)
2007 *Substrate influence in creole genesis.* [Special issue]. *Journal of Pidgin and Creole Languages* 22(1).
Mintz, Sydeney
1971 The socio-historical background to pidginization and creolization. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 481–496. Cambridge: Cambridge University Press.

Mufwene, Salikoko
2000 Creolization is a social, not a structural, process. In Ingrid Neumann-Holzschuh & Edgar Schneider (eds.), *Degrees of restructuring in creole languages*, 65–83. Amsterdam: John Benjamins.
Mufwene, Salikoko S.
2001 *The ecology of language evolution*. Cambridge: Cambridge University Press.
Mühleisen, Susanne & Bettina Migge (eds.)
2005 *Politeness and face in Caribbean creoles*. Amsterdam: John Benjamins.
Muysken, Pieter & Norval Smith (eds.)
2016 *Surviving the Middle Passage: The West Africa-Surinam Sprachbund*. Berlin: Walter de Gruyter.
Plag, Ingo
2008 Creoles as interlangues. Syntaxtic structures. *Journal of Pidgin and Creole languages* 23(2). 307–328.
Reinecke, John
1971 Tay Boi: Notes on the Pidgin French of Vietnam. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 47–56. Cambridge: Cambridge University Press.
Roberts, Sarah J.
2000 Nativization and the genesis of Hawaiian Creole. In John McWhorter (ed.), *Language change and language contact in pidgins and creoles*, 257–300. Amsterdam: John Benjamins.
Sankoff, Gillian & Penelope Brown
1976 The origins of syntax in discourse: A case study of Tok Pisin relatives. *Language* 52(3). 631–636.
Seuren, Pieter
1998 *Western linguistics: An historical introduction*. Oxford: Blackwell.
Siegel, Jeff
2008 *The emergence of pidgin and creole languages*. Oxford: Oxford University Press.
Singler, John Victor
1996 Theories of creole genesis, sociohistorical considerations, and the evaluation of evidence: The case of Haitian Creole and the relexification hypothesis. *Journal of Pidgin and Creole Languages* 11(2). 185–230.
Singler, John Victor
2008 The sociohistorical context of creole genesis. In Silvia Kouwenberg & John V. Singler (eds.), *The handbook of pidgin and creole studies*, 332–358. Malden: Wiley-Blackwell.
Smith, Norval
2008 Creole Phonology. In Silvia Kouwenberg & John V. Singler (eds.), *The handbook of pidgin and creole studies*, 98–129. Malden, MA: Wiley-Blackwell.
Thomason, Sarah G. & Terrence Kaufman
1988 *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
Tosco, Mauro & Stefano Manfredi
2013 Pidgins and creoles. In Jonathan Owens (ed.), *The Oxford handbook of Arabic linguistics*, 495–519. Oxford: Oxford University Press.
Trudgill, Peter
2009 Sociolinguistic typology and complexification. In Geoffrey Sampson, David Gill & Peter Trudgill (eds.), *Language complexity as an evolving variable*, 98–109. Oxford: Oxford University Press.
Veenstra, Tonjes
2008 Creole genesis: The impact of the language bioprogram hypothesis. In Silvia Kouwenberg & John V. Singler (eds.), *The handbook of pidgin and creole studies*, 219–241. Malden: Wiley-Blackwell.

Versteegh, Kees
2008 Non-Indo-European pidgins and creoles. In Silvia Kouwenberg & John V. Singler (eds.), *The handbook of pidgin and creole studies*, 158–186. Malden: Wiley-Blackwell.
Winford, Donald
2003 *An introduction to contact linguistics*. Malden, MA: Blackwell.

Bettina Migge, Dublin (Ireland)

10. Varieties in contact

1. Introduction
2. Language contact
3. Dialect contact
4. A borderline case
5. Summary
6. References

1. Introduction

Speakers of different varieties come into contact on a regular basis in a number of diverse communicative situations, and it is generally accepted that contact is one of the main triggers of language change (Lucas 2015). This chapter discusses contact between (speakers of) both typologically different and structurally similar (and mutually intelligible) varieties, focusing on contact-induced change in different types of contact situation. Traditionally, contact between typologically different varieties has been discussed separately from contact between structurally similar and mutually intelligible varieties, and I too consider 'language contact' and 'dialect contact' individually, presenting some central topics within each framework. I also draw parallels between the two frameworks, and discuss whether a more global approach to analyzing varieties in contact would benefit contact linguistics: a discipline covering a disparate range of phenomena characteristic of bilingual and multilingual communities that are difficult to place within a unified structure (Winford 2007).

Research on varieties in contact has dealt primarily with contact between distinct language systems (Siegel 1985: 357), within the tradition associated with Weinreich (1953). Linguists working in this field have often concentrated on the linguistic outcomes of language contact in stable and diglossic settings; more recently, scholars have focused also on synchronic variation in bilingual communities (e.g. Poplack 1980), and the role of social factors in language contact has also been considered. While the traditional approach to language contact is grounded in the study of purely linguistic factors, Sankoff (2001) points out that the outcomes of language contact are determined by both "internal" and "external" factors; Thomason and Kaufman (1988) have argued that social (and not linguistic) factors are the primary determinants of contact-induced change, but their view has been widely contested. Contact between structurally similar and mutually intelligible varieties has not been researched in equal measure. Traditional dialectology was preoccupied with mapping peripheral localized forms in linguistically homogenous

https://doi.org/10.1515/9783110435351-010

speech communities, and mixed varieties were for a long time neglected. From the 1960s, Labovian variationist research marked a turning point in the study of dialects, with the focus shifting to language variation in heterogeneous urban communities. Trudgill (1986) established dialect contact as an independent line of enquiry within the variationist paradigm; linguists working within dialect contact typically study synchronic variation and use variationist methods to describe and explain variation in situations of second dialect acquisition, new dialect formation, and regional dialect levelling.

Thomason (2008: 47) states that it is normal for "closely related systems" to be studied using "dialectological methods", while typologically different systems are studied "as languages in contact". Kerswill and Williams (2002: 82) view dialect contact as "typologically different from language contact because it does not involve speakers learning a new language, either wholly (giving rise to varying degrees of bilingualism) or in a restricted sense (typically resulting in lexical borrowing, but without any other changes)" but instead "items can be mixed apparently at will and with minimal loss of intelligibility". More recently, Hickey noted that the difference between language and dialect contact is "more one of degree than kind" (2013: 5), thus suggesting a rapprochement between the two frameworks. Thomason (2008) also states that some contact situations constitute "borderline cases" that would benefit from a combination of methodological approaches from both frameworks. This is true because 'languages' and 'dialects' are not discrete entities with clearly defined boundaries; any differences between them are often more socio-political than linguistic. By drawing a distinction between 'language contact' and 'dialect contact', we hit upon a similar problem to that which arises in attempting to distinguish between 'languages' and 'dialects': a binary division assumes that 'language contact' yields one set of outcomes and 'dialect contact' another, and that contact scenarios fall neatly within one or the other framework.

2. Language contact

2.1. Borrowing

Thomason and Kaufman (1988: 37) define borrowing as "the incorporation of foreign features into a group's native language by speakers of that language: the native language is maintained but is changed by the addition of the incorporated features" (see also Chapter 13). While borrowing may be mutual in contact situations, there is usually asymmetry that stems from the prestige of the varieties in contact. For example, contact between the Romans and the Germanic tribes led to quite substantial borrowing of Latinisms into Germanic, while few Germanic words entered into Latin; Chinese has had a considerable impact on Korean and Japanese vocabulary, but borrowing in the opposite direction has been insignificant by contrast. There was also extensive lexical borrowing from German to Czech during German colonization of Czech lands in the 13th and 14th centuries, with Germanisms affecting a number of formal and informal domains (Newerkla 2017). Some of these borrowings have survived (*klášter* 'monastery', *knedlík* 'dumpling'), others have been replaced (*rathaus* by *radnice* 'town hall', *kšaft* by *závěť* 'will [testament]') and several underwent stylistic reallocation, with the German word occupying the lower register (and in some cases assuming a more specific semantic

function) and the domestic word the higher register (*špitál* >< *nemocnice* 'hospital', *flaška* >< *láhev* 'bottle').

Intensive, ongoing contact between individual speakers is not a necessary condition for lexical borrowing. There has been no significant cultural contact between large groups of Japanese and English speakers (Hoffer 2002), yet Anglicisms are said to comprise 7.3 % of the Japanese vocabulary (Matras 2009). Motivations for borrowing are: (1) gaps in the structural inventory of the recipient language, and (2) source variety prestige (Matras 2009). In (1), source variety items absent in the recipient variety are borrowed; common examples are cultural loans that accompany new cultural associations (e.g. food items like *pretzel* and *bruschetta*). In (2), source variety items are borrowed despite there being a parallel structure in the recipient variety. Sometimes the borrowing has a nuance of meaning or connotation: French *au contraire* is more ironic (when used by native speakers of English) than English *on the contrary*, for instance. 'Prestige' is not analogous with 'more standard'. British English has borrowed *pal* 'friend (colloquial)' and *chav* 'rogue youth' from Angloromani, a variety that is neither prestigious nor powerful in Britain; these borrowings carried 'covert prestige' within the population in which they became established and later spread (Matras 2009: 151).

Some scholars propose that structural borrowing takes place only between typologically similar varieties (e.g. Winford 2005); others (notably Thomason and Kaufman 1988) argue that any linguistic feature can be transferred from one language to another, regardless of typological similarity, and consider social factors the primary determinants of contact-induced change. Unlike lexemes, structural phenomena are not unequivocally identifiable to a particular source, or an association is less transparent. Take, for example, contact between Czech and German. At the lexical level, it is hard to imagine a counter-argument to *štrúdl* being a German borrowing (*Strudel*): the word-initial consonant cluster *št-* is not observed in domestic words, word-initial [ʃt] is characteristic of German, as is represented orthographically in the Czech adapted form, and there is a considerable stock of similar borrowings (*štempl* 'stamp' < *Stempel*, *štafle* 'stepladder' < *Staffel*). At the structural level, the relationship between similar phonological or morphosyntactic structures is more opaque. Czech has resultative constructions, typical of German, made up of the verb *mít* 'to have' and a passive participle. Drinka (2017: 379) writes that resultative constructions in West Slavonic languages have "clearly developed from long-standing contact with German"; others, however, reject a contact hypothesis and argue that these constructions have developed independently. Moreover, structural changes occur in cases of language shift and are activated by processes other than borrowing.

Linguists have put forward convincing hypotheses for structural borrowing between typologically different varieties, and Seifart (2015) argues that the borrowability of certain structural elements is not affected by structural similarity. Several case studies are given in Matras and Sakel (2007); Sakel (2010) found evidence of native languages in Latin America borrowing grammatical forms, such as coordinators, discourse markers, and subordinators, from Spanish and Portuguese. Czech has arguably borrowed structural elements from German: the recipient passive made up of auxiliary *dostat* 'to get' and a passive participle (*dostat zaplaceno* 'to get paid'), inverted numbers, like *dvaadvacet* 'twenty two' (= *dva* 'two' + *a* 'and' + *dvacet* 'twenty'; lit.: 'two-and-twenty'; cf. Germ. *zweiundzwanzig*) alongside domestic *dvacet dva*, and resultative constructions (see above) made up of *mít* 'to have' and a passive participle (*mám uvařeno* 'I have done the cooking'; cf. German constructions with *haben* and a past participle [*Ich habe gekocht*]).

The influence of German is observed in other typologically different varieties (e.g. Slovak, Polish, and Hungarian) spoken in Central Europe, where cross-linguistic convergence has taken place and the emergence of a German-Slavonic-Hungarian *sprachbund* has been proposed (Thomas 2008). Quantitatively measuring similarities between the modal system of German and that of the other languages in the area in a study of contact-induced grammaticalization, Hansen (2005) observed that structural similarity correlated with geographic proximity and that modal verbs in the West Slavonic languages are considerably influenced by German. Borrowing is clear at the lexical level; for example, German *müssen* 'must' has been borrowed by several languages: *muset* (Czech), *musiet'* (Slovak), *musieć* (Polish), and *muszáj* (Hungarian).

2.2. Imposition

Structural changes activated during language shift are brought about by a process labelled 'imposition' (van Coetsem 1988), or 'interference through shift' (Thomason and Kaufman 1988). The distinction between borrowing and imposition is not clear-cut, nor is it always possible to identify which change is caused by which process, as both commonly occur in the same contact situation (Treffers-Daller 1999). In borrowing, the agents of change are speakers of the 'recipient' variety, who adopt features of the source variety, while in imposition the agents of change are speakers of the 'source' variety, who *impose* their features on the recipient variety during language shift. In imposition, phonological transfer is more common than lexical transfer; as Thomason and Kaufman (1988: 38) note, imposition "begins with sounds and syntax".

Berger (2016) suggests that diphthongization ([o] > [ou] and [ɛ] > [ɛj]) and monophthongization ([ie] > [iː] and [uo] to [uː]) processes taking place in Czech in the 14th and 15th centuries were strikingly similar to those observed in Middle High German. As Czech had gained ground as an administrative language in the late 14th century, there was a shift from German to Czech in many towns that had a German majority, and Berger puts forward an imposition-based hypothesis for these changes. Herold (1997) offers an interesting imposition-based hypothesis for the low back merger of the vowels /ɔː/ and /ɒ/ (*cot-caught*) in eastern Pennsylvania. Though a defining feature of dialects spoken in western Pennsylvania, this merger was not observed in eastern Pennsylvania, except in certain towns in which it was widespread. Herold proposes that these towns had seen a sudden massive influx of immigrants − mainly speakers of Polish and other Slavonic languages who did not distinguish between /ɔː/ and /ɒ/ − during the mining boom at the turn of the 20th century, and in shifting to English, they triggered the vowel merger. Possible examples of grammatical imposition include the durative use of *stay* in Maltese English (e.g. *I don't have the patience to stay reading a magazine*), under the influence of Maltese *qagad* 'to stay' (Bonnici 2010), and the habitual aspect (made up of *do* + *be*: *He does be in his office every morning.*), the immediate perfective (*after* + present participle: *He's after writing the letter.*), and the resultative perfective (*have* + past participle: *He has the job done*.) in Irish English under the influence of similar constructions in Irish (Hickey 1995).

Treffers-Daller (1999) analyzed borrowing and imposition in contact between Germanic and Romance varieties in Brussels and Strasbourg. Her data support Thomason and Kaufman's (1988) hypotheses in that contact phenomena observed in Brussels Dutch

and Alsatian spoken in Strasbourg are almost exclusively the result of lexical borrowing, while the dominant French varieties spoken in both cities display clear traces of imposition, especially at the phonological level. Word-final obstruent devoicing, a tendency towards diphthongization, and default stress at the beginning of a word in Brussels French are all listed as phenomena resulting from the influence of Dutch.

2.3. Code-switching

Code-switching, typical of the speech of bilinguals, denotes alternation between two language varieties and is rule-governed (Poplack 1980). A distinction is drawn between *insertional* and *alternational* code-switching. Huang and Milroy (1995: 36–37) describe the difference as follows: in alternational code-switching, "two (or more) languages are used in turn in the same clause or discourse", while in insertional code-switching, "a word or phrase from one language is put into the grammatical framework of a clause or discourse made up of another language" (1995: 36–37). In alternational switching bilinguals possess a high level of proficiency in both languages, while this is not necessarily true of insertional code-switching (see also Chapter 14).

As borrowings, insertional code-switches can compensate for a lexical gap, be used for special effect, or have a nuance of meaning absent in the parallel structure. Matras (2009: 102–107) gives an example of code-switching in the speech of a Romani speaker from Hamburg, who inserts German words for institutions and institutional procedures (e.g. *Meldeamt* 'registration office', *Sterbeurkunde* 'death certificate') because (1) there are no parallel structures in the speaker's native variety, and (2) use of the original word "evokes associations with the original setting and allows the speaker to import the image of that setting directly into the context of the ongoing conversation" (2009: 107). Understandably, insertional code-switching has been likened to borrowing (Matras 2009: 110–114), but, while borrowings are the product of a diachronic process, an insertional code-switch is used spontaneously and need not gain currency within the wider community (i.e. insertional code-switching is a precursor to borrowing, but need not result in borrowing).

3. Dialect contact

3.1. New dialect formation and reallocation

New dialect formation is defined by Kerswill and Williams (2005: 1023–1024) as "the emergence of distinctive, new language varieties following the migration of people speaking mutually intelligible dialects to linguistically near-'virgin' territory" (e.g. 'new towns', such as Milton Keynes [Kerswill and Williams 2005] and former British colonies [Trudgill 2004]). It has also been studied in speech communities whose demographic profile changed rapidly due to mass migration (Honeybone 2007). New varieties arise out of contact between two or more competing varieties and subsequent koineization, consisting of the levelling of linguistically marked and marginal forms, the emergence of intermediate (interdialect) forms, and simplification. First-generation contact is char-

acterized by considerable inter-individual variation that gives rise to interdialect features that did not occur in any of the source varieties, but which emerge from contact and mixing between them (Trudgill 1986: 62). Once these interdialect features crystallize and are adopted as native variants, a new dialect is born.

The time it takes for a stable variety to emerge depends on the heterogeneity of the speech community. In the Czech Sudeten lands of northern Bohemia, migrants were predominantly from other parts of Bohemia, a region characterized by linguistic homogeneity, and Common Czech, a supralocal koine spoken throughout Bohemia, emerged as the stable variety by the second generation. By contrast, new dialect formation in the Šumperk and Bruntál areas of northern Moravia, populated by migrants from nearby towns and villages, and from other parts of the Czech and Slovak Republics, re-emigrants from Czech-speaking communities in Russia and Romania, Germans, and later, Romany Gypsies and migrants from Macedonia and Greece, was more complicated; there, a stable variety emerged in the third generation (Kloferová 2000). Honeybone (2007) looks at new dialect formation in the 19th century in Liverpool, where mass migration led to the emergence of features that make Liverpool linguistically distinct. Contact phenomena include a merger to [ɛː] in the NURSE and SQUARE lexical sets (*square* [skwɛː], *nurse* [nɛːs]), and 'Liverpool lenition', a process by which plosives are realized as affricates and fricatives in certain environments (e.g. *crime* realized as [kxɾaːm]).

Occasionally, two or more forms survive after koineization and assume different socio-stylistic or structural functions in the new dialect. Migration to Norwich from surrounding rural areas triggered change in the urban dialect (Britain and Trudgill 2005), with most rural, peripheral forms being levelled and a single variant surviving after koineization. However, three variants survived in a subset of the GOOSE lexical set (*room*, *broom, groom*), [uː], [ʊ], and [ʉː], with each variant being refunctionalized and marking social status: [uː] (high), [ʊ] (middle), and [ʉː] (low). This is an example of *socio-stylistic reallocation*. Australian English provides us with an example of *structural reallocation*. In the BATH lexical set, both 'northern' /æ/ and 'southern' /aː/ were present after koineization, but they have been reallocated to different phonological environments: although their distribution varies regionally, /æ/ is usually observed before nasal clusters (e.g. *dance*) and /aː/ before other phonological contexts (Britain and Trudgill 2005).

3.2. Second-dialect acquisition (SDA) and regional dialect levelling

Migration brings about changes to both the recipient and the source varieties. Urban varieties show most innovation, given that migration to large cities has been most intensive. As the varieties in contact are mutually intelligible, social factors override linguistic factors in determining the direction and the degree of language change. Migrants typically accommodate to their new speech community (Giles 1973: 90) to (1) gain their interlocutors' social approval, (2) ensure mutual intelligibility, and (3) maintain positive social identities. Under favorable extra-linguistic conditions, short-term linguistic modifications may trigger permanent changes to a speaker's idiolect (Trudgill 1986: 40).

SDA (see Siegel 2010) is normally incomplete in the first generation due to interference from the source variety and an incomplete knowledge of the recipient variety. Migrants mix variants of the source and recipient varieties, use recipient variety forms inconsistently, and use 'interdialect' forms that are found neither in the source nor in the

recipient variety (Trudgill 1986). Wilson (2010) identified all types of Trudgill's partial accommodation, including code mixing between Standard Czech and Common Czech forms in adjectival desinences found to be impermissible among native speakers (e.g. *to bude docela takové* (SC) *dobrý* (CC) 'that will be kind of good'). Erratic oscillation between SC and CC forms was common and is different from the rule-governed switching between source and recipient variety identified within the language contact framework. Linguistically very complex variants may not be acquired even by migrants before the critical age. An oft-cited example is acquisition of Philadelphia 'short *a*': /æ/ is realized as either [æ] or [eə] (e.g. *can* 'to be able to' is pronounced [kæn], but *can* 'metal container' is pronounced [eə]) and has complex distributional rules that inhibit SDA (Payne 1980).

Extralinguistic factors can hinder or accelerate SDA. While living in the USA, Trudgill (1986) avoided adopting features of the recipient variety that were "too American" (e.g. /æ/ in *dance*). In this case, non-accommodation did not impede social integration because the source form is intelligible and non-stigmatized in the recipient community; moreover, Trudgill was a short-term migrant. By contrast, Kerswill (1994: 37), in a study of accommodation in Norway, reported that speakers of Stril dialects from rural areas around Bergen who had moved to Bergen for work were told by their employers that if they did not speak "properly" they would have to leave their job. Here the motivation to accommodate was strong because non-accommodation impedes integration.

Migration affects also the recipient variety. Recipient variety speakers are exposed to new forms and mutual accommodation leads to the levelling of peripheral forms in all varieties, potentially triggering *regional dialect levelling*: the loss of localized linguistic forms at the expense of linguistic features with a wider areal distribution. Regional dialect levelling consists of *levelling* and *geographical diffusion* (Kerswill 2002). *Levelling* denotes a loss of irregularity and is closely connected to mutual accommodation. Geographical, social, and occupational mobility has triggered the elimination of many peripheral linguistic features because in contact situations, highly localized linguistic forms are marked, and are thus likely to be eliminated, while supralocal forms typically prevail. *Geographical diffusion* denotes the spread of linguistic forms from a dominant linguistic center to a wider region. The diffused forms may be the result of earlier contact-induced levelling in the urban center from which they spread and may spread first via other, smaller, urban centers to the countryside.

Regional dialect levelling in England is widespread: a clear example is the linguistic homogenization in the South East (Kerswill 2002). Kerswill argues that certain linguistic forms are even spreading throughout the country: for example, intervocalic *t*-glottaling (i.e. the replacement of /t/ by [ʔ] between vowels [*better*]) and *th*-fronting (the merger of /θ/ and /f/ [*thing*], and of /ð/ and /v/ medially [*brother*] and finally [*smooth*]). Other examples of geographical diffusion include (1) the spread of pan-northern /eː/ in the FACE set (e.g. *cake*, *safe*, *late*) at the expense of local [ɪə] in Newcastle and Durham, and (2) the spread of Common Czech, a koine that developed in the 14th and 15th centuries in Central Bohemia, to other areas of Bohemia, with localized forms such as /ɛj/ in the locative/dative singular of adjectives and pronouns (*tej dobrej* 'that good'), typical of southwest Bohemian dialects, being replaced by Common Czech /iː/ (*tý dobrý*).

4. A borderline case

Some cases of contact do not fit neatly within either the language contact or dialect contract framework. An example is contact between Czech and Slovak. These two 'languages' are structurally very similar and by and large mutually intelligible (Nábělková 2007).

As in cases of language contact, there has been mutual lexical borrowing. Historically, Czech has had a major influence on the development of Slovak, and Czech has borrowed many words from Slovak, in particular during the 18th and 19th centuries. Lexical borrowing between the two languages is very productive also in the modern era. Slovak, for example, has borrowed *venčiť* 'to walk' (e.g. *venčiť psa* 'to walk the dog') from Czech *venčit* (*venčit psa*) owing to language economy, as the domestic equivalent is a multi-word expression (*chodiť von so* [*psom*]). Structural similarity allows for the easy borrowing of morpho-syntactic structures between Czech and Slovak (Musilová 2011). Corpus data show a rise in the use of the genitive case in Slovak after pronouns like *niečo* 'something' and *nič* 'nothing' (e.g. *niečo iného* 'something else') in place of domestic nominative/accusative (*niečo iné*) under the influence of Czech (*něco jiného*). In Czech the prepositional phrase *k ránu* 'in the early hours (of the morning)' (*k* + the dative form of *ráno* 'morning') has lost ground to *nad ránem* (*nad* + the instrumental form) under the influence of Slovak *nad ránom*.

Phenomena typical of dialect contact are also observed. In Czech-Slovak communication, speakers can stick to their native language, shift to the other language, or mix linguistic features of both languages. It seems to come down to personal preference whether Slovaks living in the Czech Republic shift to Czech or not (Nábělková 2007: 59). Those who do switch manifest incomplete SDA. Code mixing is also observed, as in the following utterance of Pavol Habera, a Slovak panel member on the TV talent show *Česko-slovenská SuperStar* (from Musilová 2011):

> … má množstvo **fanyniek**. Už **kecám blbo**. Nebudem **si delať prču** z **takových** vecí. … To bolo **šílené**! **Přímo prúser jak** mraky!
>
> '… he has loads of female fans. I'm already being silly. I shouldn't make fun of things like this. … It was crazy! What a cock-up!'

The sections highlighted in bold are Bohemianisms. Nouns (*fanyniek*, *prču*), adjectives (*takových*, *šílené*), verbs (*kecám, delať*), adverbs (*blbo*, *přímo*), and conjunctions (*jak*) are easily borrowed into Slovak; similar inflectional rules allow for the direct copying of some features (e.g. the genitive plural adjective *takových* [Slovak = *takých*]).

5. Summary

At a very general level, the outcomes of contact between all language varieties, typologically different or structurally similar, are the same: contact leads to (1) change affecting one, some, or all varieties in contact, (2) the formation of new varieties, or (3) the loss of existing varieties.

In migration to other speech communities, regardless of typological similarity, first-generation (adult) migrants' idiolects undergo change; there is inter- and intra-individual variation and acquisition of the recipient variety is usually incomplete. There is interference from the source variety and migrants usually hit a plateau in recipient variety proficiency. Attrition of the source variety might arise as a consequence of recipient variety acquisition. In contact between structurally similar varieties, migrants' idiolects can become estranged from the source variety to the point that migrants are mistaken as outsiders in their native speech community, but do not attain sufficient proficiency in the second variety to pass as natives of the recipient community (Wilson 2010). The speech of second-generation migrants is usually indistinguishable from that of members of the recipient community whose parents are from that community, but there are exceptions, especially when migrants move to highly segregated societies or to areas that are heavily populated by other migrants from the same source culture.

There are clear similarities in the way that certain new 'languages' and 'dialects' develop. Take, for example, *pidginization* and *koineization*. Siegel (1985: 370–376) writes that (1) both processes involve contact between different varieties, and that a new variety emerges through this contact and subsequent mixture; (2) both koines and pidgins display simplification and reduction; (3) the development of a *pre-pidgin* into a *creole* is strikingly similar to that of a *pre-koine* into a *nativized koine*, with both varieties undergoing expansion of function and form. There are, of course, differences. Other than the fact that the varieties involved in the formation of a pidgin are usually typologically different, while koines are formed through the mixing of structurally similar and/or mutually intelligible varieties, pidgins undergo more extensive reduction and simplification than koines, and pidginization is a rapid process, whereas koineization "is a slow, gradual process which requires continued contact and integration among the speakers of the different varieties" (Siegel 1985: 376).

There are also notable differences. In contact between typologically different varieties, lexical borrowing is common, while there may be little, if any, structural borrowing after centuries of intensive cultural contact (unless language shift occurs). Conversely, structural borrowing, including the borrowing of inflectional morphology, considered particularly resistant to transfer, commonly occurs in first-generation contact between 'dialects' (e.g. Kerswill 1994; Wilson 2010) as well as between structurally similar 'languages', like Czech and Slovak. Structured code-switching is a frequent outcome of contact between typologically different varieties, but not between structurally similar varieties; in contact between varieties with a similar morphosyntactic structure, frequent, and more haphazard, mixing of items is observed.

There is an important difference within the two frameworks with regard to the consideration of social factors in the outcomes of contact. In dialect contact, social factors are at the forefront of analysis; this is logical in that speakers of mutually intelligible varieties need not change the way they speak in contact situations, and convergence through accommodation and SDA is thus socially motivated. The converse is true in language contact: social factors have received comparatively little attention. Scholars may not have recourse to the necessary extralinguistic data to investigate the social mechanisms underlying diachronic contact-induced change. Discipline territoriality may also play a role in that it has simply been the norm for those working on language contact to study the *linguistic* constraints on borrowing and code-switching. Thomason and Kaufman (1988) mark a shift of focus in this regard, although later empirical works (e.g. Treffers-

Daller 1999) reject their hypotheses that there are no absolute structural constraints on borrowing and still consider linguistic factors the primary determinants of contact-induced change. Sociolinguistic work that has been carried out since the 1960s shows that language variation and change is conditioned by the complex interplay of language-internal and extralinguistic factors, and this observation no doubt extends to language contact situations.

Current research into varieties in contact embraces both traditions and their methodologies, and this is clearly a step in the right direction. Contact between two typologically very different languages will inevitably yield some different outcomes to contact between two neighboring dialects, but a binary division of contact situations seems counter-productive; it encourages fragmentation and hinders a holistic approach to contact linguistics. The outcomes of contact are best seen along a continuum according to typological difference and mutual intelligibility, and it makes sense for contact linguistics to study contact between all varieties under one umbrella to achieve a better understanding of the mechanisms underlying contact-induced change.

6. References

Berger, Tilman
2016 The convergence of Czech and German between the years 900 and 1500. In Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder & Achim Rabus (eds.), *Congruence in contact-induced language change: Language families, typological resemblance and perceived similarity*, 189–198. Berlin: Walter de Gruyter.

Bonnici, Lisa
2010 *Variation in Maltese English: The interplay of the local and the global in an emerging postcolonial variety*. Berkeley: University of California. PhD dissertation.

Britain, David & Peter Trudgill
1999 Migration, new-dialect formation and sociolinguistic refunctionalisation: Reallocation as an outcome of dialect contact. *Transactions of the Philological Society* 97(2). 245–256.

Britain, David & Peter Trudgill
2005 New dialect formation and contact-induced reallocation: Three case studies from the Fens. *International Journal of English Studies* 5(1). 183–209.

Drinka, Bridget
2017 *Language contact in Europe: The periphrastic perfect through history*. Cambridge: Cambridge University Press.

Giles, Howard
1973 Accent mobility: A model and some data. *Anthropological Linguistics* 15. 87–105.

Hansen, Björn
2005 How to measure areal convergence: A case study of contact-induced grammaticalization in the German-Hungarian-Slavonic contact area. In Björn Hansen & Petr Karlík (eds.), *Modality in Slavonic languages*, 219–237. Munich: Sagner.

Haugen, Einar
1950 The analysis of linguistic borrowing. *Language* 26(2). 211–231.

Herold, Ruth
1997 Solving the actuation problem: Merger and immigration in eastern Pennsylvania. *Language Variation and Change* 9(2). 165–189.

Hickey, Raymond
1995 An assessment of language contact in the development of Irish English. In Jacek Fisiak (ed.), *Linguistic change under contact conditions*, 109–130. Berlin: Mouton de Gruyter.

Hickey, Raymond
2013 Language contact: Reconsideration and reassessment. In Raymond Hickey (ed.), *The handbook of language contact*, 1–28. Oxford: Blackwell.
Hoffer, Bates L.
2002 Language borrowing and language diffusion: An overview. *Intercultural Communication Studies* 11(4). 1–37.
Honeybone, Patrick
2007 New-dialect formation in nineteenth century Liverpool: A brief history of Scouse. In Anthony Grant & Clive Grey (eds.), *The Mersey sound: Liverpool's language, people and places*, 106–140. Liverpool: Open House Press.
Huang, Guowen & Lesley Milroy
1995 Language preference and structures of code-switching. In David Graddol & Stephen Thomas (eds.), *Language in a changing Europe*, 35–46. Clevedon: Multilingual Matters.
Kerswill, Paul
1994 *Dialects converging: Rural speech in urban Norway*. Oxford: Clarendon Press.
Kerswill, Paul
2002 Models of linguistic change and diffusion: New evidence from dialect levelling in British English. *Reading Working Papers in Linguistics* 6. 187–216.
Kerswill, Paul & Ann Williams
2002 "Salience" as an explanatory factor in language change: Evidence from dialect levelling in urban England. In Mari C. Jones & Edith Esch (eds.), *Language change: The interplay of internal, external and extra-linguistic factors*, 81–110. Berlin: Mouton de Gruyter.
Kerswill, Paul & Ann Williams
2005 New towns and koineization: Linguistic and social correlates. *Linguistics* 43(5). 1023–1048.
Kloferová, Stanislava
2000 *Mluva v severomoravském pohraničí* [Everyday speech in the border regions of northern Moravia]. Brno: Masarykova univerzita.
Lucas, Christopher
2015 Contact-induced language change. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–537. New York: Routledge.
Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.
Matras, Yaron & Jeanette Sakel (eds.)
2007 *Grammatical borrowing in cross-linguistic perspective*. Berlin: Mouton de Gruyter.
Musilová, Květoslava
2011 Slovakismy v současné češtině (sociolingvistický průzkum mezi vysokoškoláky) [Slovakisms in modern Czech (a sociolinguistic investigation of university students)]. In Martin Ološtiak, Martina Ivanová & Daniela Slančová (eds.), *Vidy jazyka a jazykovedy*, 389–396. Prešov: Philosophical Faculty of Prešov University.
Myers-Scotton, Carol
1992 Comparing codeswitching and borrowing. *Journal of Multilingual and Multicultural Development* 13(1–2). 19–39.
Nábělková, Mira
2007 Closely-related languages in contact: Czech, Slovak, "Czechoslovak". *International Journal of the Sociology of Language* 183. 53–73.
Newerkla, Stefan M.
2017 Germanismy v českém lexiku [Germanisms in Czech]. In Petr Karlík, Marek Nekula & Jana Pleskalová (eds.), *Nový encyklopedický slovník češtiny* [New Encyclopedic Dictionary of Czech], 584–588. Prague: Lidové noviny.

Payne, Arvilla
1980 Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In William Labov (ed.), *Locating language in time and space*, 143–178. New York: Academic Press.
Poplack, Shana
1980 Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics* 18(7–8). 581–618.
Sakel, Jeanette
2010 Grammatical borrowing from Spanish/Portuguese in some native languages of Latin America. *Language Typology and Universals* 36(1). 65–78.
Sankoff, Gillian
2001 Linguistic outcomes of language contact. In Peter Trudgill, J. K. Chambers & Natalie Schilling-Estes (eds.), *Handbook of sociolinguistics*, 638–668. Oxford: Blackwell.
Seifart, Frank
2015 Does structural-typological similarity affect borrowability? A quantitative study on affix borrowing. *Language Dynamics and Change* 5(1). 92–113.
Siegel, Jeff
1985 Koines and koineization. *Language in Society* 14(3). 357–378.
Siegel, Jeff
2010 *Second dialect acquisition*. Cambridge: Cambridge University Press.
Thomas, George
2008 Exploring the parameters of a Central European Sprachbund. *Canadian Slavonic Papers* 50(1–2). 123–131.
Thomason, Sarah G.
2008 Social and linguistic factors as predictors of contact-induced change. *Journal of Language Contact* 2(1). 42–56.
Thomason, Sarah G. & Terrence Kaufman
1988 *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
Treffers-Daller, Jeanine
1999 Borrowing and shift-induced interference: Contrasting patterns in French-Germanic contact in Brussels and Strasbourg. *Bilingualism: Language and Cognition* 2(1). 1–22.
Trudgill, Peter
1986 *Dialects in contact*. Oxford: Blackwell.
Trudgill, Peter
2004 *New-dialect formation: The inevitability of colonial Englishes*. Edinburgh: Edinburgh University Press.
van Coetsem, Frans
1988 *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
Weinreich, Uriel
1953 *Languages in contact: Findings and problems*. The Hague: Mouton.
Wilson, James
2010 *Moravians in Prague: A sociolinguistic study of dialect contact in the Czech Republic*. Frankfurt: Peter Lang.
Winford, Donald
2005 Contact-induced changes: Classification and processes. *Diachronica* 22(2). 373–427.
Winford, Donald
2007 Some issues in the study of language contact. *Journal of Language Contact* 1(1). 22–40.

James Wilson, Leeds (UK)

11. Language contact and constructed languages

1. Introduction
2. Terminology
3. *A priori* and *a posteriori* language creation
4. Change in artificial languages
5. Conclusion
6. References

1. Introduction

People who construct languages − whether called Esperanto (Zamenhof 1887) or Dothraki (Peterson 2011−2016) − typically do so because they are somehow dissatisfied with the set of existing languages: those may either be considered inadequate instruments for thought or communication, or too difficult to learn; they may further not fit the imaginary world of a fiction writer. Interlinguistics, the field that studies such languages (see Schubert [1989] for an overview), makes the common distinction between *a posteriori* and *a priori* constructed languages. The former are built on the model of existing languages: Peano's (1903) Latine sine flexione, for instance, which, as its name suggests, is basically a version of Latin without the inflectional morphology. The latter are constructed 'from scratch', such as Lojban (Cowan 1997), which is supposed to provide a purely logical way of expressing thoughts.

The *a priori-a posteriori* distinction should not be seen as a clear-cut binary opposition, but rather as a continuum between these two extremes. On the one hand, *a posteriori* languages will always display elements of willful design, based on some *a priori* idea of how languages can be improved. On the other hand, it is probably not difficult to show that *a priori* language creators are influenced by the languages they already know, and that this influence is similar to the impact of language contact.

Artificial or constructed languages are interesting for any scholar of language contact by their very nature. They present extreme cases of contact: extreme in the level of consciousness that is involved in their planning, and extreme in that we can typically point to an originator or a committee of originators. In the (rare) case where people adopt artificial languages in their everyday life, for instance, as a family language, this language will undergo even further contact, because there are no communities in which one can live one's entire life speaking an artificial language. There will always be an 'outside world' in which other languages are used. Speakers of artificial languages will therefore always be at least bilingual and the other 'natural' language(s) in their life will always be dominant.

In this chapter, I discuss both parts of language contact in constructed languages. I concentrate mostly on languages from the 19th and 20th century. Those of the 17th century were mostly 'philosophical' and not necessarily meant to resemble existing languages in any way; they eventually led to notation systems in logic and mathematics. Languages of the 21st century are usually designed for use in fictional worlds and often meant to embody the imaginary character of those settings (see Peterson [2015] for a nice introduction to the application of linguistic insights in this kind of language design). The most interesting contact phenomena we find, I think, in languages that were designed for human and international use, and that has mostly been a preoccupation of the 19th and 20th century.

https://doi.org/10.1515/9783110435351-011

After a terminological discussion in section 2, section 3 addresses the relevance of the *a priori* vs. *a posteriori* dimension in our understanding of language contact, as well as the way language projects are influenced by existing languages. The development of artificial languages when they become part of a community is at the heart of section 4, with a clear emphasis on Esperanto as the artificial language with the longest history and the widest application. Section 5 summarizes some implications of these findings for the field of contact linguistics.

2. Terminology

All artificial languages differ from 'natural' languages in one shared characteristic: we can identify one or more human beings who have put some conscious effort into their 'creation'. These languages did not exist before that effort, although it is not immediately clear at what point (if ever) they actually became 'languages'. While the conscious act of creation may not have been enough, it would typically involve setting up both some vocabulary − usually with translation of a few words into some other, already existing language − and some rules of grammar. It is a matter of debate whether those systems can already be considered to constitute a language in their own right. I suppose most scholars would say that more is needed, although there could be an argument as to what exactly that would entail.

One would typically expect one or more people to use the language for some purpose and this active use (and the accompanying acts of the speakers) are at least as important for 'creating' the languages as the initial steps set by the initiator. As soon as this happens and these languages start functioning more or less 'normally', however, there is no longer any reason to expect them to develop differently from other languages.

The terminological question now is twofold: what do we call the first step, and how do we identify the languages that are eventually the product of this change? So far, I have used the terms 'artificial languages' (Comrie 1996) and 'constructed languages'; additional labels used in the literature include 'planned languages' (with *planning* as the relevant act), 'invented languages' and 'synthetic languages'.

Constructed language seems to be the most popular term among (English-speaking) scholars (definitely distinct from the German literature, for instance, where *Plansprache* is a common term; Blanke 1985), whereas artificial language may be the more popular reference among the general public. In linguistics, constructed language is also used for experimental work with made-up stimuli (Miller 1958; Moro 2016). A recent collection of scholarly work (Adams 2011a) uses invented languages. Planned language is possibly a somewhat wider term, as standard languages are all likely to have undergone some level of planning. One interesting case in point is Modern Hebrew, which has been 'planned into existence' to some extent (Berdichevsky 2014; see Romaine [2011] for revitalized languages more generally): it would not have existed in its present form, were it not for the active involvement of a certain number of intellectuals. Modern Hebrew can therefore be called a 'planned language', but the terms 'constructed' or 'artificial' seem less suitable for this language.

In this paper, I restrict myself to the use of the terms 'artificial' and 'constructed language' because they are the most frequent, and because I want to *exclude* languages

like Hebrew and Indonesian from consideration here. It is useful, however, to realize that the boundaries between the languages discussed here, and other (standardized) languages, are not precisely delineated. A central claim of this paper will therefore be that not all aspects of 'constructed languages' are indeed carefully crafted, so that there is quite some space for influence of other languages (Adams 2011b).

In this context, similar concerns can be raised about the term *language*, particularly when debating the question at which point a constructed object can rightfully be called a 'language'. Blanke (1989) proposed a useful taxonomy in this respect, with a three-way distinction between 'projects', 'semi-languages', and 'languages'. Projects are basically objects which never grew out of the initial grammars and lexicons written by their constructors. They can grow into semi-languages and eventually even into languages, once they gradually meet an increasing list of criteria, such as being used in written communication, appearing in books and journals, or having specialized vocabulary pertaining to a number of topics and domains. Languages are distinguished from semi-languages by having native speakers. According to Blanke (1989: 72), Esperanto is the only full language in the set of constructed objects. Given that Blanke was a leading figure in the Esperanto movement, however, one might suspect that this judgement was biased, but at the same time, Esperanto has undeniably been most widely applied in a wide variety of linguistic contexts (see also Pereltsvaig [2016] for an overview of linguistic literature on Esperanto).

A final distinction needs to be made between constructed languages, secret languages, and/or language games (Bausani 1974). The latter are more clearly parasitic on an existing language, in the simplest case by replacing single sounds with other sounds in a systematic way, or by distorting the phonology in some other ways, to give but two possible examples. The boundaries between these intriguing systems and constructed languages proper seem easier to draw in practice than in theory (one knows a secret language when seeing it, but it remains difficult to define); they fall beyond the scope of this chapter, however.

3. *A priori* and *a posteriori* language creation

Since it seems impossible to create something as complex as a human language from scratch, language makers by necessity depend on existing languages, be it subconsciously or consciously. Subconscious influence typically occurs at the phonetic, grammatical, semantic, and/or pragmatic levels, whereas more conscious influence appears at the lexical level. For instance, Interlingua was designed by a committee − for some time under the direction of the well-known French structuralist linguist André Martinet − which used very specific rules for the inclusion of a form in the morpheme stock of the language. Using a precise method to count in how many European languages a form occurred, this form would only be chosen if the count transcended an established threshold (Martinet 1946). Furthermore, the committee also defined rules for deriving the Interlingua grammar: the smallest common denominator was to be chosen, and only rules which were present in the Romance languages and English were supposed to be operative. Yet, since we know that there is no full description of any language in the world, it is equally

impossible to write a complete *proscription* of a language. As a matter of fact, the overt rules are often defined only negatively, in the sense that the new language distinguishes itself through the absence of certain features (e.g. morphological case) from other languages known to the author or to the intended readership. A careful scholar may therefore presumably extract regularities from an Interlingua corpus that are not explicitly described in the official grammar. These hidden rules could, for instance, be borrowed from the native language of the speaker.

We can take the following passage on adjective placement from the grammar of Interlingua (Gode and Blair 1951: 12), as an example:

> Adjectives placed next to a noun can either *precede or follow.* The latter position is more frequent and hence normal. Adjectives preceding a noun tend to suggest that what they express is an essential feature of the noun concept and not merely a feature distinguishing the present representative of the noun concept from others.
>
> *le lingua international e le linguas national*
> 'the international language and the national languages'
>
> *Su integre vita esseva ric in viages longe e breve*
> 'His whole life was rich in long and short trips'
>
> Long adjectives should rarely be allowed to precede a noun. Brief adjectives like *bon, alte, parve, grande, belle, breve, longe*, etc., may precede merely for rhythm's sake or as a matter of personal preference. The two possible positions of the adjective cannot ever express two distinct meanings. Numeral adjectives (cardinals as well as ordinals) precede the noun they qualify.

The tone of this passage is descriptive; it bases its recommendations regarding what to say on what is 'more frequent' and 'normal'. Considering there had been very few publications in Interlingua at the time of its publication, it is unclear how 'frequency' could even be measured. However, the tentative rules which are formulated here − there is a specialized interpretation for prenominal adjectives, there is more freedom for shorter adjectives to precede the noun, and there is a 'personal preference' − resemble those of (modern) Romance languages. Wim Jansen (personal communication) points out that, for example, *povre* 'poor' seems to differentiate between prenominal use (where it means 'pitiful') and postnominal use (where it means 'without sufficient material support'). A similar distinction is made in Italian, for instance. It is as if the authors (both English-speaking Americans) refer the reader to linguistic intuitions for Romance, which make it unnecessary to be more explicit, as the reader is supposed to share those intuitions.

Many grammars of language projects share the common practice of using the terminology of neutral description, as if the language is already there. They especially tend to take many things for granted and fail to provide, or ignore, certain explanations. In the example cited above, for example, the grammar does not give any answer to the question of how to deal with cases where there is more than one adjective with a noun and one does not want to use the conjunction *e*. This may appear to turn such grammars into a parody of 'real grammars', describing a phantasy rather than a reality. We cannot exclude, however that the authors of these grammars had a real feel about what 'sounded' usual or unusual. Moreover, statements like these in traditional grammars of 'real' natural languages may not be always based on careful counting of items in a high-quality corpus, either.

As a matter of fact, we can even observe that we find similar statements in literature proposing standardization. Standardization in some cases also means trying to change a language in such a way that it starts resembling some other language more closely, albeit in a superficial level. In the standardization of European languages, for example, the urge to copy the model of Latin would lead to the creation of tables with grammatical cases, even if some languages no longer had cases (Eco 1993).

A posteriori languages can further be subdivided into those which are based on just one language, such as Basic English (Ogden 1930) or Latino sine flexion (Peano 1903), and those which are based on some kind of typological comparison, such as most other language projects of the 19th and 20th century, such as Interlingua and also Volapük (Schleyer [1880] 1982). In the former cases, the goal is usually to 'simplify' the language by, for example, eliminating grammar that is deemed unnecessary (as *sine flexione* 'without flection' suggests), or by reducing the vocabulary (this was the main goal of Basic English: to allow for communication using only words from a list of a few hundred vocabulary items); see Gobbo (2017). In the latter case, creators aim to make a language that represents the 'greatest common denominator' of all languages in the sample. For instance, Schleyer wanted to make the words of Volapük easy to pronounce for people around the world, and he therefore eliminated the distinction between *l* and *r*. At the same time, he preserved umlauted letters which, like in his native German, stood for front rounded vowels. These vowels are not exactly widespread in the languages of the world. The umlauted graphemes further received quite a bit of criticism and may have accounted, according to some historians, for the language's eventual lack of success (Garvía 2015).

The question of how to 'blend' languages in this way has occasionally attracted the attention of professional linguists as well. Trubetzkoy (1939), for instance, discusses the question of which phonemes should be used in an ideal language for international communication, with a view towards maintaining sufficient contrasts between sounds, while at the same time avoiding major problems for any speaker.

All artificial languages, including heavily *a priori* systems, are obviously influenced by existing 'natural' languages to some extent, if only because they are designed to be *languages.* As such, they will typically have some concept of a word and of a way of arranging those words (i.e. syntax), as one cannot imagine a language being otherwise. This was definitely true even in the case of 17th century creators of 'philosophical' languages, such as Leibniz and Wilkins; their creations were not intended to merely resemble existing 'natural' languages, but rather to be more perfect tools for thinking (Maat 2004). These languages use their own systems for word formation, basing every individual element on some idea of rationality and logic, rather than on what exists in language. Yet, as far as I am able to tell, notions similar to words and sentences are always there, just as formal logic derived from those 17th century attempts always is. One could argue, admittedly, whether such divisions are *derived* from natural languages, or whether those languages just reflect something which is fundamentally logical. Revealing the underlying structure which was available in all languages, without the historically developed and 'unnecessary' additions, appears to have been the intent of many of the language creators of this period.

Furthermore, in view of the thousands of decisions that have to be made when one tries to create a language *ab ovo,* language creators will inevitably turn to linguistic traits that seem familiar, or even inevitable, to them. The most dominant languages for Lejzer

Zamenhof, the creator of Esperanto, presumably were Russian and Yiddish (Schor 2016), but he was also familiar at least with Hebrew, Polish, and German. Being the 19th century language lover that he was, most of Zamenhof's conscious efforts were directed at building a morphological system in which all nouns end in an *-o*, all adjectives in an *-a*, all infinitives of verbs in an *-i*, etc., which makes it easy to convert words from one category into another (*bono* is 'goodness', *bona* 'good', *boni* 'being good'). He also cared about the etymological origin of the word stock, which is mostly Romance, probably because Latin was a former international lingua franca and French was the most important language for international trade and diplomacy of the 19th century. However, some of the words are also of Slavic (*barakti* 'to struggle') or Germanic (*hundo* 'dog') origin.

Zamenhof published a list of 16 'grammatical rules' which have very little to say about either the phonology or the syntax of Esperanto (one rule says 'the definite determiner is *la*'). There is probably a relation between this marginal attention to both grammatical properties and the fact that the consonant inventory of Esperanto is exactly the same as that which Zamenhof described in his grammar of Yiddish (Piron 1984). As a matter of fact, his Esperanto orthography even uses more or less the same symbols for the same sounds as found in his earlier proposal for a Latinate script for Yiddish. The syntax of Esperanto likewise resembles that of Russian and/or Yiddish in several ways (for instance, its basic word order) and can probably be accounted for as subconscious influence. The influence of Yiddish on Esperanto has been pointed out by several scholars (Biró 2004; Gold 1980; Golden 1982; Lindstedt 2009; Piron 1984; Schor 2016), just as the influence of Russian has been explored in Kolker and Piron (2015). There is very little reason to assume that this was a conscious choice on Zamenhof's part, however. The creator most probably saw no need to take any decision on these matters and just chose the kind of linguistic solutions that felt most natural to him from his practical experience. It has been argued, as a matter of fact, that even the consciously designed morphology had its origins in existing language. Zamenhof himself reported that the idea for a very regular derivational agglutinative morphology came to him as a child when he observed that the words *shvejcarskaja* 'porter's lodge' and *konditorskaja* 'candy shop' both contained the locational suffix *-skaja* (Zamenhof 1948: 133). In turn, the idea that a word category (noun/adjective/verb/adverb) can have a phonological exponence might be borrowed from Hebrew (Berdichevsky 1986, 2007).

An interesting twist to language contact in language design is provided by those constructed languages that were designed to be as *unnatural*, or as different from existing languages, as possible in certain dimensions. Examples of this are Klingon (Okrand 1985; Okrand et al. 2011) and Spocanian (Tweehuysen 1982). The latter language, imagined to be used in an imaginary but realistic country Spocania, expresses tense by word order (past tense is VSO, present tense SVO, future tense SOV), a property that does not exist in any other language. It is probably significant that such languages are usually made for fantasy worlds. They typically are constructed by people with some linguistic training who are aware of the linguistic rules they are breaking. The languages in question could be claimed to have undergone a *negative* influence of language contact, in that they were consciously steered in the opposite direction from the languages they were in contact with.

4. Change in artificial languages

While the majority of artificial languages never leave the drawing board of their makers, some do make it to a small community. In very rare occasions (as for Esperanto), they can even become one of the native languages of multilingual speakers. In such cases, structural influence from the other languages used by members of the community can occur. The likely need to elaborate the word stock to include new concepts in the language will equally involve some concept of 'borrowing' (see Chapter 13).

Esperanto could be said to come closest to being a 'natural' language, as it is the artificial language which is both the most well-known, and the one with the vastest range of application. It was adopted by a diasporic community of a few tens of thousands of speakers (Nielsen 2016), including a group of about one thousand native speakers, according to an estimate by Lindstedt (1996). It seems that those numbers have stayed relatively stable over the ensuing decades.

Several studies have appeared over the years addressing diachronic changes in Esperanto. These include lexical innovation through borrowing, both for new concepts (*apo* 'app', for example, in the sphere of new technology) and for aspects of daily life of Esperanto-speaking families which were not foreseen by Zamenhof (*butikumi* 'go shopping'). When discussing lexical change over time in Esperanto, Philippe (1991) points out that new Esperanto words were modeled as loan translations on elements from different dominant languages in different time periods (see also Gregor 1974). For instance, in the first period (in roughly the last decade of the 20th century), loans were generally modeled on either Russian or German. The word for 'security' was *sen-danĝer-ec-o* (I added hyphens for morphological transparency), which can be glossed as 'without-danger-property-noun', modeled on Russian *bezopasnost'*. The word for 'singular' was *unu-nombr-o* (one-number-noun), modeled on German *Einzahl.* In some cases, the model is Yiddish rather than German, as in *lern-o-libr-o* (learn-noun-book-noun, 'textbook'), which does not directly correspond to German *Lehrbuch* ('teach-book'), but more directly to Yiddish *lernbukh* (Lindstedt 2009).

Around the beginning of the 20th century, these words were replaced in common practice by *sekur-ec-o* (safe-property-noun) and *singular-o*, respectively. This was a period in which the center of the movement was to be found in France. Other loan translations from this period can also be brought back to this language. For instance, Philippe (1991) points to *flav-bek-ul-o* 'inexperienced young person' ('yellow-mouth-person-noun'). The model for this word was French *bec-jaune* ('mouth-yellow'), but the Esperanto word has a different order, and contains a personifying suffix in order to conform to Esperanto (i.e. Yiddish, German, or Russian) morphology. Contemporary loans are instead built on English models: *korp-o-konstru-o* ('body-noun-build-noun', 'bodybuilding'), *sub-urb-o* ('under-city-noun', 'suburb').

Lindstedt (2009) discusses how there also was a semantic shift in the early period of Esperanto, away from the Russian origins. In early Esperanto texts, *naĝi* ('to swim') could also be used for 'ships', like Russian *plyt'* and *plavat'*, but this is no longer possible.

In an experimental study, Dankova (1997, 2009) shows how contemporary Esperanto speakers with different (Slavic vs. Romance) linguistic backgrounds use the verbal system of the language in different ways to express temporal and aspectual details when telling a story; their respective choices align with the systems used in their native lan-

guage. She points out that Zamenhof did not establish a 'norm' for this (although his own use of the system was unsurprisingly more on the Slavic side). Interestingly, Jansen (2007) has very different findings for word order in Esperanto. Although the Esperanto 'ideology' permits all possible word orders (e.g. SVO, SOV, VSO, VOS, OVS, OSV for the order of subject, verb, object, and AN, NA for noun, adjective), Jansen found that in actual fact, speakers in experiments had an overwhelming preference for a very small range of orders (SVO and AN). This preference was also reflected in several corpora of written Esperanto consulted by Jansen.

It is not clear why word order tends to become fixed across speakers of very different languages, whereas tense/aspect stays more flexible; the fact that the first is more clearly located on the surface (and therefore easier to observe) than the latter may be relevant. It is also not entirely clear why the preference went to SVO and AN. Other word orders did occur in the corpus and in experimental settings, for instance, to express information structure. Jansen (2007) did already find a preference for the basic orders in a corpus of Zamenhof's own writings, however, as early as his first texts in Esperanto, before there even were any other speakers of the language. Maybe this is (just) another example of Zamenhof's dominant languages subconsciously influencing his grammatical choices. Lindstedt (2009: 12) mentions that "perhaps the most deep-seated Slavonic feature" in Esperanto syntax is the tendency to place manner adverbs before verbs as in 'mi bone scias tion', 'I know that well'

For an alternative analysis, one could approach Esperanto as a form of pidgin or creole (Heil 1999; Parkvall 2008, 2010). These word orders might then be the result of forces similar to those which typically cause contact languages to have a certain 'default' order. In that case, even the very early 'community', consisting of only Zamenhof, would have been subject to such forces.

The 'native speakers' of Esperanto are of course also of interest in this connection. This is a small group of about one thousand people spread around the world. They are hardly a community, none of them is monolingual, and they always have another dominant language. Furthermore, almost all of them are first-generation native speakers, although some second and maybe third generation speakers have also been reported. Some work has been done on their first language acquisition (Bergen 2001; Versteegh 1993), but in my feeling, this is hardly enough to draw good conclusions about the process. An interesting point is put forward by Pereltsvaig (2016), who mentions that, according to Corsetti (1996), in most families it is the father who speaks Esperanto to the child. It could be interesting to look into the precise implications of this predominant 'paternal' language transmission.

5. Conclusion

Any kind of language creation is inevitably influenced by existing languages in one way or another. This seems true even for the more extreme cases: languages that do not even pretend to be used in everyday human communication, like programming languages, the 'language' of logic, or the languages that have been designed for science fiction movies. It is particularly true, however, for those languages that aim to compete with existing languages in the domain of (international) communication.

It is important to point out that designing a truly complete language appears to be as equally impossible as describing one in its entirety. If all grammars leak, so too do all prescriptions. If such projects are to grow into real languages, the 'holes' have to be filled, and elements from the existing languages known to the speakers of the new project will fill those holes. This transforms artificial languages into interesting laboratories for finding out how both language and language contact, in particular, work. For this reason, it is unfortunate that constructed languages hitherto have not received a lot of attention, although this seems to be gradually changing in the scholarly literature.

Acknowledgments

This paper has gained from insightful comments by Federico Gobbo, Wim Jansen, the editors, and two anonymous reviewers of this volume.

6. References

Adams, Michael (ed.)
2011a *From Elvish to Klingon: Exploring invented languages*. Oxford: Oxford University Press.

Adams, Michael
2011b The spectrum of invention. In Michael Adams (ed.), *From Elvish to Klingon: Exploring invented languages*, 1–17. Oxford: Oxford University Press.

Bausani, Alessandro
1974 *Le lingue inventate.* Milano: Astrolabio Ubaldini.

Berdichevsky, Norman
1986 Zamenhof and Esperanto. *Ariel: A Review of Arts and Sciences in Israel* 64. 58–71.

Berdichevsky, Norman
2007 Why Esperanto is different. *New English Review*. http://www.newenglishreview.org/Norman_Berdichevsky/Why_Esperanto_Is_Different/ (accessed on 7 March 2018).

Berdichevsky, Norman
2014 *Modern Hebrew: The past and future of a revitalized language*. Jefferson, NC: McFarland & Company.

Bergen, Benjamin K.
2001 Nativization processes in L1 Esperanto. *Journal of Child Language* 28(3). 575–595.

Biró, Tamás
2004 Weak interactions: Yiddish influence in Hungarian, Esperanto and Modern Hebrew. In Dicky Gilbers, Maartje Schreuder & Nienke Knevel (eds.), *On the boundaries of phonology and phonetics*, 123–145. Groningen: University of Groningen.

Blanke, Detlev
1985 *Internationale Plansprachen: Eine Einführung.* Berlin (Ost): Akademie-Verlag.

Blanke, Detlev
1989 Planned languages: A survey of some of the main problems. In Klaus Schubert (ed.), *Interlinguistics: Aspects of the science of planned languages*, 63–77. Berlin & New York: Mouton de Gruyter.

Comrie, Bernard
1996 Natural and artificial international languages: A typologist's assessment. *Journal of Universal Language* 1(1). 35–55.

Corsetti, Renato
1996 A mother tongue spoken mainly by fathers. *Language Problems & Language Planning* 20(3). 263–273.
Cowan, John Woldemar
1997 *The Lojban reference grammar.* Fairfax: The Logical Language Group.
Dankova, Natalia
1997 Temporalité en espéranto: Étude du transfert. PhD Dissertation, Paris 8.
Dankova, Natalia
2009 Temporality in spoken Esperanto. *Esperantologio/ Esperanto Studies* 4. 43–60.
Eco, Umberto
1993 *La ricerca della lingua perfetta nella cultura europea*. Roma-Bari: Laterza.
Garvía, Roberto
2015 *Esperanto and its rivals: The struggle for an international language.* Philadelphia: University of Pennsylvania Press.
Gobbo, Federico
2017 Are planned languages less complex than natural languages? *Language Sciences* 60. 36–52.
Gode, Alexander & Hugh E. Blair
1951 *Interlingua: a grammar of the international language*. New York: Storm Publishers.
Gold, David L.
1980 Towards a study of possible Yiddish and Hebrew influence on Esperanto. In Istvan Szerdahélyi (ed.), *Miscellaneous interlinguistics*, 300–367. Budapest: Tankönyvkiadó.
Golden, Bernard
1982 The supposed Yiddish origin of the Esperanto morpheme *edz*. *Jewish Language Review* 2. 21–33.
Gregor, Douglas Bartlett
1974 La plej fruaj kunmetaĵoj en Esperanto. *Scienca Revuo* 25(2–3). 81–102.
Heil, Anett
1999 *Grammatische Reduktion in Frankokreolsprachen und Plansprachen*. Frankfurt/Main: Peter Lang.
Jansen, Willem Hendrik
2007 *Woordvolgorde in het Esperanto: Normen, taalgebruik en universalia.* Utrecht: LOT.
Kolker, Boris & Claude Piron
2015 *La kontribuo de la rusa al Esperanto: Influo de la jida sur Esperanton: Esperanto: ĉu lingvo okcidenta.* Paris: Monda asembleo socia.
Lindstedt, Jouko
1996 Kiom? http://web.archive.org/web/20060924124158/http://www.rano.org/nombro.html (accessed on 7 March 2018).
Lindstedt, Jouko
2009 Esperanto: An East European contact language? In Christian Voß & Alicja Nagórko (eds.), *Die Europäizität der Slawia oder die Slawizität Europas: Ein Beitrag der kultur- und sprachrelativistischen Linguistik*, 125–134. München & Berlin: Otto Sagner.
Maat, Jaap
2004 *Philosophical languages in the 17th century: Dalgarno, Wilkins, Leibniz.* Dordrecht: Kluwer.
Martinet, André
1946 La linguistique et les langues artificielles. *Word* 2(1). 37–47.
Miller, George Armitage
1958 Free recall of redundant strings of letters. *Journal of Experimental Psychology* 56(6). 485–491.
Moro, Andrea
2016 *Impossible languages.* Cambridge: MIT Press.

Nielsen, Svend
2016 Per-country rate of Esperanto speakers. *Kalkulinda. Visualisations and probabilities.* https://svendvnielsen.wordpress.com/2016/12/10/percountry-rates-of-esperanto-speakers/ (accessed on 7 March 2018).
Ogden, Charles Kay
1930 *Basic English: A general introduction with rules and grammar*. London: Paul Treber.
Okrand, Marc
1985 *The Klingon dictionary.* New York: Pocket Books.
Okrand, Marc, Michael Adams, Judith Hendriks-Hermans & Sjaak Kroon
2011 Wild and whirling words: The invention and use of Klingon. In Michael Adams (ed.), *From Elvish to Klingon: Exploring invented languages*, 111–134. Oxford: Oxford University Press.
Parkvall, Mikael
2008 The simplicity of creoles in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, 265–285. Amsterdam: John Benjamins.
Parkvall, Mikael
2010 How European is Esperanto? A typological study. *Language Problems & Language Planning* 34(1). 63–79.
Peano, Giuseppe
1903 De Latino sine flexione: Lingua auxiliare internationale. *Revista de Mathematica (Revue de Mathématiques)* VIII. 74–83.
Pereltsvaig, Asya
2016 State of the art: Esperanto linguistics. http://www.esperantic.org/en/research/state-of-the-art/state-of-the-art-esperanto-linguistics-2/ (accessed on 7 March 2018).
Peterson, David J.
2011–2016 *Dothraki: A language of fire and blood.* http://www.dothraki.com/ (accessed on 22 March 2017).
Peterson, David J.
2015 *The art of language invention: From horse-lords to dark elves, the words behind world-building*. New York: Penguin.
Philippe, Benoît
1991 *Sprachwandel bei einer Plansprache am Beispiel des Esperanto.* Konstanz: Hartung-Gorre Verlag.
Piron, Claude
1984 Contributions à l'étude des apports du yiddish à l'espéranto. *Jewish Language Review* 4. 15–29.
Romaine, Suzanne
2011 Revitalized languages as invented languages. In Michael Adams (ed.), *From Elvish to Klingon: Exploring invented languages*, 185–226. Oxford: Oxford University Press.
Schleyer, Johann Martin
1982 [1880] *Volapük: Die Weltsprache: Entwurf einer Universalsprache für alle Gebildete der ganzen Welt.* Hildesheim: Olms. [The 1982 edition was edited by Reinhard Haupenthal.]
Schor, Esther
2016 *Bridge of words: Esperanto and the dream of a universal language*. New York: Metropolitan Books.
Schubert, Klaus (ed.)
1989 *Interlinguistics: Aspects of the science of planned languages*. Berlin & New York: Mouton de Gruyter.
Trubetzkoy, Nikolai S.
1939 Wie soll das Lautsystem einer künstlichen internationalen Hilfssprache beschaffen sein? *Travaux du Cercle Linguistique de Prague* 8. 5–21.

Tweehuysen, Rolandt
1982 *Uit in Spokanië nooit weg: Een reisverslag*. Utrecht: Gralkrich, Pelger & Mardant.
Versteegh, Kees
1993 Esperanto as a first language: Language acquisition with a restricted input. *Linguistics* 31(3). 539–555.
Zamenhof, Ludwik Lejzer
1887 *International language.* Warsaw: Chaim Kelter. [Appeared under the pseudonym D-ro Esperanto.]
Zamenhof, Ludwik Lejzer
1948 *Leteroj de Zamenhof.* Paris: SAT. [Edited by Gaston Waringhien.]

Marc van Oostendorp, Amsterdam & Nijmegen (The Netherlands)

II. Language contact and the individual

12. Pragmatics

1. Pragmatics of language contact
2. Conceptualizing language contact
3. New directions
4. Challenges for future research
5. References

1. Pragmatics of language contact

This chapter aims to depict the pragmatics of language contact as a field that is characterized both by intradisciplinary diversity and interdisciplinary convergence. Pragmatics is most often defined as studying language in use (e.g. communicative functions) and the contexts in which it is used; that is, the study of the adaptability of language to social and cultural ecologies (Verschueren 1987). Originally emerging as a branch of semiotics (Morris 1938), and forming a disciplinary triad with syntax and semantics, pragmatics might rather be considered as offering a particular perspective on many of the same objects that syntacticians, semanticists, psycholinguists, phonologists, and/or morphologists approach in their research (Verschueren 1987). As the literature referred to in this chapter will suggest, by foregrounding language use in context with reference to language contact, pragmatics has a close, dialectical relationship with sociolinguistics and linguistic anthropology/ethnography on the one hand, and with psycholinguistics on the other. Current pragmatics research further confirms that the epistemological distinction between linguistic-philosophical pragmatics (e.g. interlanguage pragmatics, cross-linguistic pragmatics), sociocultural-interactional pragmatics (e.g. sociopragmatics), as well as the more recently emerging intercultural pragmatics (e.g. intercultural politeness) set out by Horn and Kecskes (2013), also holds true for the pragmatics of language contact. Horn and Kecskes (2013: 353) define these traditions as follows:

> Linguistic-philosophical pragmatics seeks to investigate speaker meaning within an utterance-based framework focusing mainly on linguistic constraints on language use. Socio-cultural interactional pragmatics maintains that pragmatics should include research into social and cultural constraints on language use as well. The link between classical philosophically-oriented pragmatics and research in intercultural and inter-language communication has led to the development of intercultural pragmatics, focusing on the roles and functions of language and communication within a world-wide communication network. Intercultural pragmatics attempts to combine the two traditions into one explanatory system that focuses special attention on characteristics of intercultural interaction.

Thus, a major difference between linguistic-philosophical pragmatics and other approaches is the examination of not just the individual utterance, but discourse and interaction.

This chapter is structured into four sections. After this section, section 2 will take issue with the ontological question of how language contact in diverse strands of prag-

https://doi.org/10.1515/9783110435351-012

matics research has been conceptualized, and trace the foundational influences of these strands. Section 3 will discuss emerging directions in pragmatics scholarship. The chapter concludes with some of the current challenges in the pragmatics of language contact. While aiming to be inclusive throughout, sociocultural-interactional and intercultural epistemologies are foregrounded in the latter two sections, as necessitated by the selective nature of the discussion.

2. Conceptualizing language contact

Schiffrin's (1996) discussion on the pragmatics of language contact focused on three discursive features − codeswitching, contextualization cues, and speech acts − used by socially and culturally heterogeneous groups. Although these aspects remain relevant more than three decades on, and would in themselves be deserving of a thorough review, researchers nowadays contribute to the pragmatics of language contact by responding to multifarious issues and from diverse paradigms. Thus, some of the conceptualizations of 'language' and of 'contact' emerging from recent research are discussed briefly in this section. We take a bottom-up approach to the question of how language contact is treated empirically in pragmatics, and how it falls within certain 'classic' traditions. While the research has been grouped into four broad perspectives, these are not meant to be bounded categories, and some of the studies approach language contact in intersecting ways.

In a significant number of studies, *named languages are brought into contact by researchers*, in the sense that features of communication in language X are contrasted with features of communication in language Y. This is especially the case in subfields such as intercultural pragmatics or cross-linguistic pragmatics. For example, Lorenzo-Dus and Bou-Franch (2013) studied a corpus of impromptu, monolingual e-mails by speakers of either Peninsular Spanish or British English. Their results show complex, fluctuating patterns in levels of (in)formality and (in)directness, which they related to cross-cultural variation in the way that different sociopragmatic principles found expression in computer-mediated communication. Zufferey, Mak, and Sanders (2015) investigated the acquisition of objective and subjective causal relations in monolingual French- and monolingual Dutch-speaking children from age five to eight on the same comprehension task. Their findings show that French- and Dutch-speaking children have a similar ability to handle the linguistic features studied, thereby refuting arguments of linguistic determinism on acquisition. In both of these studies, the languages/varieties in question (Peninsular Spanish and British English; French and Dutch) were brought into contact by researchers through the design of the studies, rather than being a feature of a particular individual's or social collective's repertoire. Many of the studies that conceptualize language contact in this way have their foundations in contrastive linguistics, represented in the early work of scholars such as Lado (1957), and also find illustrious precedents in some of Clyne's (1987) early work on written texts.

In a similarly large number of studies, *language contact is studied as a phenomenon at the level of individual speakers' repertoires*. This is the case in much research on intercultural and interlanguage pragmatics, including intercultural politeness. For example, Safont-Jordà (2013) conducted a longitudinal study of the early stages of trilingual pragmatic development, focusing on requests in Catalan, Spanish, and English. The re-

search highlighted how the introduction of a third language (i.e. English) in a young boy's already bilingual (i.e. Catalan and Spanish) repertoire prompted the use of conventionally indirect request forms in the three languages. Gassner (2012) compared how L1 Australian English speakers, and L2 English speakers who had migrated to Australia from different South-American, European, and/or Asian countries, used the word *thing* in an oral corpus of job interviews, concluding that the L2 speakers were more vague in their speech. Parvaresh (2015) responded to this study by using a corpus of speech by L1 Persian and L2 English speakers to show how vagueness is equally a feature of L1 and L2 repertoires. Mugford (2013) used questionnaires and interviews with female French and German learners of Spanish in Mexico to examine their use of *güey*, a seemingly impolite word used in an inoffensive way, finding that they failed to use it for the whole spectrum of its functions. Dewaele (2016) demonstrated that speakers who acquire English later in life generally overestimate the offensiveness of certain emotion-laden words, are less sure about their meaning, and tend to use them less frequently as compared to L1 English speakers. However, the more contact they have with monolinguals, the more the L2 speakers' use converges with that of L1 speakers. Most of these studies focus on how users of multiple languages use the pragmatic features of them, hinting at how transfer takes place across the languages in their repertoires. A significant body of research understanding language contact in this way has a basis in classic work on interlanguage (Selinker 1972; Weinreich 1953).

In a third set of studies, especially those within sociopragmatics, *language contact is explored in terms of communicative encounters in everyday contexts of linguistic and cultural diversity.* This research often has foundations in early work on code choice and codeswitching (Gumperz 1982; Myers-Scotton 1993; Poplack 1980), in microsociological work on aspects such as politeness (Goffman 1959), and in discourse analytic approaches to intercultural pragmatics (Tannen 2005), among other influences. For example, Ifukor (2011) studied language use in Nigerian online forums, claiming that linguistic diversity is particularly of interest in computer-mediated communication due to the dominance of English in this medium. The researcher showed how the informality of internet discussions allows significant insights into 'linguistic eclecticism', language contact, and codeswitching to be gathered, as characteristics of vernacular talk and internet multilingualism in unregulated spaces. Albirini (2014) looked at code-switching in spontaneous speech, at weddings, and in interviews, by speakers of three varieties of Arabic in Syria: a standard variety, an urban variety, and a Bedouin variety, drawing on Bourdieu's theory of social capital. The author concludes that use of one variety or another is sociopragmatic, in that switching can serve functions such as indexing identities, helping to develop intra-tribal connections, and facilitating access to workplace relationships. Androutsopoulos (2014) studied entextualization and linguistic repertoires in social networking, and specifically showed how young Greek-German Facebook users mobilized languages (e.g. German, Greek, and English) to entextualize different events. He argues that language choice in this online space was connected to sociodemographic characteristics of the users, as well as transnational trajectories (e.g. travelling to Greece). The public space of social networking thus provided participants with novel opportunities to display language use in relation to mobility, be it a new language encountered or one hidden in a user's repertoire. These studies, therefore, all involve naturally-occurring interactions in different social settings in

which the use of more than one language or variety emerges amongst speakers with heterogeneous repertoires.

Lastly, in a smaller set of studies, *language contact is conceptualized in terms of encounters between modes, media, genres, registers, etc.*, thereby offering an expanded perspective on what language is by incorporating recent advances in the study of multimodality (Kress 2009; Kress and van Leeuwen 2001). For example, Tseng (2016) took a sociopragmatic and discourse analytic approach to examine advertisements in Taiwan, which are predominantly in Chinese, although some English words might be used in the product title, text, and/or picture(s). The analysis showed how multiple meanings are created in a multilingual context: using more than one language enriches meaning-making and contributes to consolidating the creative force of the product discourse. Sambre and Feyaerts (2017), in their study on a trumpet master class, highlight the interplay between gesture, speech, playing instruments, and the presence of material objects for the achievement of musical interpretation. Their research points clearly at the combination of modes by taking into account the embodied experience of subjects in relation to their surroundings. Chiluwa and Ajiboye (2016) adopted a discourse-pragmatic approach to examine the various ways youths construct themselves, their group identities, and their environment and socio-economic aspirations by using T-shirt messages and slogans at two educational institutions in Nigeria. The messages on the T-shirts, which are mainly in Nigerian English, combine features of spoken and written communication, including text message language, and some Nigerian Pidgin. Similar to the previous set of studies, this research deals with language use in contexts of diversity, concentrating on how meaning is created in relation to localities shaped by the social presence of varied linguistic repertoires.

3. New directions

The origins of language contact as a field of inquiry are often traced to scholars such as Haugen (1953), Weinreich (1953), Ferguson (1959), or Fishman (1967). While pioneering an emerging discipline, these early scholars also implicitly supported a view of languages as used in separate domains. The different languages known by multilingual individuals were often described as employed for doing different things, with different interlocutors, keeping less dominant linguistic resources for the more private domains of life, and prestigious, standard languages for more public ones (i.e. diglossia). Such conceptualizations have been challenged in recent research, so much so that scholars such as Jaspers and Madsen (2016: 236) have even proposed that "what we know today as a standard language is an idea that has depended on assiduously separating language from nature and society before associating it with civilisation, progress, and later with national, supposedly organic communities". However, these earlier ideas fit well with the understanding at the time in structural linguistics (e.g. Lado 1957), which continues to persist in much scholarship today: that languages could be compared in order to predict sources of cross-linguistic transfer by learners in order to avoid them, and to keep languages bounded.

A number of studies have started to engage with pragmatics in a way that sees language contact differently. In particular, some research has taken up the call voiced by

an increasing number of scholars for studying language use from a perspective that considers that "named languages are social constructs and not lexical or structural ones" (Otheguy, García, and Reid 2015: 287). Indeed languages are increasingly conceptualized in terms of "entities without names, as sets of lexical and structural features that make up an individual's repertoire and are deployed to enable communication" (Otheguy, García, and Reid 2015: 286). These are ideas that are not new, having previously been raised by scholars such as Alvarez-Caccamo (1998), Gafaranga (2000), or Heller (2002), and are traceable in the work of Hymes (1972); however, they have found new impetus in recent years. This shift is in keeping with much psycholinguistics research that has stressed the relevance of re-thinking the conventional separation between different linguistic subsystems in the brain in favor of an integrated system of multi-competence (Cook and Li Wei 2016).

Numerous terms have been proposed recently in disciplines concerned with the social use of language to describe communicative practices that draw on whole repertoires in an expansive way. These include 'languaging' (Becker 1995), 'plurilingual practices' (Lüdi and Py 2009), 'translanguaging' (García 2009), 'codemeshing' (Canagarajah 2011), 'metrolingualism' (Otsuji and Pennycook 2010), and 'polylingualism' (Jørgensen 2008), as well as the multi-competence approach already mentioned, first put forward by Cook (1991). Many of these terms at present are beginning to find representation in pragmatics research. Of them, translanguaging, or "the deployment of a speaker's full linguistic repertoire without regard for watchful adherence to the socially and politically defined boundaries of named (and usually national and state) languages" (Otheguy, García, and Reid 2015: 283) is gaining significant acceptance by researchers from an array of disciplines concerned with the social use of language. The notion includes codeswitching, the focus of much pragmatics research until now, but it also goes beyond it. It refers not only to fluid practices that go between and beyond national and state languages, but also to those that go between and beyond visual, aural, embodied, and spatial modalities. Interestingly, although translanguaging is just starting to find its place in pragmatics, Li Wei's (2011) seminal article on translanguaging was published in the *Journal of Pragmatics,* thereby speaking directly to research on the pragmatics of language contact, and suggesting the breadth of theories that feed into pragmatics.

Recent research in a Polish shop in East London (Zhu Hua, Li Wei, and Lyons 2017) has shown the applicability of translanguaging to the analysis of interactions in space, as well as the complex practices that are integral parts of the communication in the shop itself. These include gestures, gazes, and movement, but also the display and arrangements of signs and objects. Similarly, Lou (2017) was able to integrate the analysis of verbal interaction with a range of other modes through what she called a 'geosemiotic' analysis, showing how different Hong Kong markets display, and allow individuals to use, various forms of semiotic resources. Similar interests have also driven Pennycook and Otsuji (2017) to conceptualize semiotic assemblages, and other scholars (Izadi 2017; Karrebæk 2017; Kusters 2017; Williams 2017) to approach markets so that semiotic resources, rather than linguistic units, are brought to the fore as agentive in communication. This is in line with the view put forward by Li Wei (2018: 26) as follows:

> We know for a fact that the labelling of languages is largely arbitrary and can be politically and ideologically charged, and there is often a close relationship between the identity of a language and the nation-state. But in everyday social interaction, language users move

> dynamically between the so-called languages, language varieties, styles, registers, and writing systems, to fulfil a variety of strategic and communicative functions. The alternation between languages, spoken, written, or signed; between language varieties; and between speech, writing, and signing, is a very common feature of human social interaction. It constructs an identity for the speaker that is different from a La identity or a Lb identity. Moreover, language users use semiotic resources, gesture, facial expression, etc., in conjunction with language to communicate with each other. From a Translanguaging perspective, asking simply which language is being used becomes an uninteresting and insignificant question.

Even if we consider recent studies that do not take these novel perspectives into account explicitly, we can see that the range of linguistic elements brought into the analysis of interactions is increasingly multi-layered. This is, for instance, the case in a study by Okada (2015), which used Conversation Analysis (CA) to analyze audio-recorded job interviews for a student assistant position in English classes at a Japanese university, part of which was conducted in the named languages, English and Japanese. The author analyzed the data in English, which included some Japanese features. Yet rather than focusing on the juxtaposition of English and Japanese in what might be defined as insertional codeswitching (Auer 1999), the author centered the analysis on how different communicative moves, including different lexical features of speakers' repertoires, contributed to building rapport through sequentially linked joke-serious responses. In Italy, Baraldi (2018) showed that a Nigerian interpreter in a migration support center not only translates and helps with mutual understanding, but also plays a major role in reducing cultural distance. Furthermore, it is the interaction between the migrant, the social worker, and the interpreter that contributes to the articulation of problems as they progressively emerge in multilingual encounters. In Australia, Hlavac (2014) studied the participation roles of a language broker and the discourse of brokering among Macedonians in Australia. He showed how language brokers assumed the roles of animators, authors, and principals, following Goffman (1981), in communicating between languages; they do not simply translate passively (animators) or put together the pieces (authors), but also state beliefs, positions, etc., using the first person (principal). Such fluid language use is arguably more complex than codeswitching or translation, and appears to require new descriptive terms. Studies such as these show how the tools of pragmatics can shed light on meaning-in-interaction when resources are managed simultaneously in context.

These resources may well go above and beyond the tools that speakers employ in face-to-face interaction. As Androutsopoulos (2013) points out, digitally-mediated interaction means that languages are coming into contact in written as well as spoken modalities like never before, although research in this area is still comparatively scarce. There are several exceptions to this in recent pragmatics research mentioned in the previous section, including Ifukor (2011) and Androutsopoulos (2014). The complexity of handling interactional data that is both multilingual and multimodal was also highlighted recently by Egbert, Yufu, and Hirataka (2016). Recent contributions to pragmatics by authors including Broth and Mondada (2013), who study video recordings of guided tours in French, Swedish, and English, or Ticca and Traverso (2017), who examine video recorded encounters in which French mediators offer assistance to migrants for social, legal, and health matters, provide methodological ways forward for handling complex embodied interactional data. The special issue of the *Journal of Pragmatics* edited by Mondada (2014) is also insightful in this regard. Social-semiotic approaches to multi-

modality, represented in recent pragmatics work by authors including Moya Guijarro (2011), also offer potential for understanding linguistic diversity as it emerges in visual modes.

4. Challenges for future research

The scenario that researchers have to handle is a complex one, where language use is not only hard to interpret because it encompasses a range of manifestations that have not been taken into account before, but also because these manifestations question the way in which we see language contact as a whole. If we follow a perspective whereby language contact is fully embedded in language use in all its forms, then the way in which we conceptualize contact when undertaking pragmatics research needs to be revisited. As we have seen in the previous sections, scholars from various traditions have approached contact differently, and have succeeded in showing that language use does vary across contexts and that communicative norms play a role in shaping the way in which communication is organized and achieves its goals. As Gafaranga (2017) argues, now that the linguistic behavior that bilinguals display is being liberated from its old, negative connotations, we need to move forward in the study of the interactional practices it involves.

Researchers from a range of disciplines within and beyond linguistics use the term 'superdiversity', coined by Vertovec (2007), to refer to the linguistic and cultural diversity of 21st century urban contexts. The term, which is not without controversy (e.g. Reyes 2014), has been useful for framing linguistic intricacies that include historical diversities, the results of demographic mobility (e.g. refugees and asylum seekers, labor migration, tourism, and/or international education), and transformations in the way we communicate brought about by digital technologies (Appadurai 1996; Blommaert 2010; Blommaert and Rampton 2011; Castells 1996). Being theorized as a form a renovation as opposed to a total reinvention (Arnaut et al. 2015), superdiversity, for linguistics, means focusing on the sheer complexity of communication as it unfolds at this historical juncture.

In pragmatics research, scale theory has been proposed as a possible model to look at the relationship between linguistic phenomena and social structures (Blommaert 2007), in a bid to investigate language while considering that different individuals have uneven access to linguistic resources, and that the way they mobilize them is influenced by higher power dynamics. These are now being reshaped, as mobility is being integrated in the very concept of context, and the distinction between context and text becomes increasingly problematic (Blommaert 2017b; Canagarajah 2017). If context is not simply something that contains and supports communication, but plays a much more active role, then indeed its reconceptualization is imperative in pragmatics research. Such a reconsideration calls for both spatial and chronological depth (e.g. Blommaert 2017a), as well as a rethinking of the ways in which language contact can be framed to account for its interdependencies with human mobility, or lack thereof.

Focusing on these complexities is indeed related to the understanding of the actual relationship between what happens at the micro-level of communication, and the larger phenomena at the societal level that have an impact on how language takes shape. In other words, continuing to recombine and problematize 'micro-' and 'macro-' aspects

(Blommaert 2017a; Lo and Park 2017) is a key challenge that the study of language faces, and that scholars will have to address in years to come. This is in line with a view of multilingualism that sees conversational dynamics in co-production with larger systems and structures, where the resources speakers use are both connected to, and possibly transcend, situations of mobility (Rampton 2011). Further taking up this challenge could effectively mean that the mobility that linguistic resources are embedded into, and the subsequent contact that occurs among them, will have to be considered as critical factors for an appreciation of their functions in society. In this sense, contact, rather than constituting a point of encounter among well-segmented resources, would be seen as part and parcel of how individuals use language as they do the work of meaning-making.

5. References

Albirini, Abdulkafi
2014 The socio-pragmatics of dialectal codeswitching by the Al-'Keidaat Bedouin speakers. *Intercultural Pragmatics* 11(1). 121–147.

Alvarez-Caccamo, Celso
1998 From 'switching code' to 'code-switching': Towards a reconceptualization of communicative codes. In Peter Auer (ed.), *Code-switching in conversation: Language, interaction, and identity*, 29–48. London: Routledge.

Androutsopoulos, Jannis
2013 Code-switching in computer-mediated communication. In Susan Herring, Dieter Stein & Tuija Virtanen (eds.), *Pragmatics of computer mediated communication*, 667–694. Berlin & Boston: De Gruyter Mouton.

Androutsopoulos, Jannis
2014 Moments of sharing: Entextualization and linguistic repertoires in social networking. *Journal of Pragmatics* 73. 4–18.

Appadurai, Arjun
1996 *Modernity at large: Cultural dimensions of globalization*. Minneapolis: University of Minnesota Press.

Arnaut, Karel, Jan Blommaert, Ben Rampton & Massimiliano Spotti
2015 Introduction: Superdiversity and sociolinguistics. *Tilburg Papers in Culture Studies* 130. 1–17.

Auer, Peter
1999 From code-switching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* 3(4). 309–332.

Baraldi, Claudio
2018 Interpreting as mediation of migrants' agency and institutional support: A case analysis. *Journal of Pragmatics* 125. 13–27.

Becker, Alton L.
1995 *Beyond translation*. Ann Arbor: University of Michigan Press.

Blommaert, Jan
2007 Sociolinguistic scales. *Intercultural Pragmatics* 4(1). 1–19.

Blommaert, Jan
2010 *The sociolinguistics of globalization*. Cambridge: Cambridge University Press.

Blommaert, Jan
2017a Chronotopes, scales and complexity in the study of language in society. In Karel Arnaut, Martha Sif Karrebæk, Massimiliano Spotti & Jan Blommaert (eds.), *Engaging super-*

diversity: Recombining spaces, times and language practices, 47–64. Bristol: Multilingual Matters.

Blommaert, Jan
2017b Commentary: Mobility, contexts, and the chronotope. *Language in Society* 46(1). 95–99.

Blommaert, Jan & Ben Rampton
2011 Language and superdiversity. *Diversities* 13(2). 1–22.

Broth, Mathias & Lorenza Mondada
2013 Walking away: The embodied achievement of activity closings in mobile interaction. *Journal of Pragmatics* 47(1). 41–58.

Canagarajah, Suresh
2011 Codemeshing in academic writing: Identifying teachable strategies of translanguaging. *The Modern Language Journal* 95(3). 401–417.

Canagarajah, Suresh
2017 Introduction. In Suresh Canagarajah (ed.), *The Routledge handbook of migration and language*, 1–28. New York: Routledge.

Castells, Manuel
1996 *The rise of the network society, the information age: Economy, society and culture*, vol. I. Malden, MA: Blackwell.

Chiluwa, Innocent & Esther Ajiboye
2016 Discursive pragmatics of T-shirt inscriptions: Constructing the self, context and social aspirations. *Pragmatics and Society* 7(3). 436–462.

Clyne, Michael
1987 Cultural differences in the organisation of academic texts. *Journal of Pragmatics* 11. 211–247.

Cook, Vivian
1991 The poverty-of-the-stimulus argument and multi-competence. *Second Language Research* 7(2). 103–117.

Cook, Vivian & Li Wei
2016 *The Cambridge handbook of linguistic multi-competence*. Cambridge: Cambridge University Press.

Dewaele, Jean-Marc
2016 Thirty shades of offensiveness: L1 and LX English users' understanding, perception, and self-reported use of negative emotion-laden words. *Journal of Pragmatics* 94. 112–127.

Egbert, Maria, Mamiko Yufu & Fumiya Hirataka
2016 An investigation of how 100 articles in the Journal of Pragmatics treat transcripts of English and non-English languages. *Journal of Pragmatics* 94. 98–111.

Ferguson, Charles F.
1959 Diglossia. *Word* 15(2). 325–340.

Fishman, Joshua A.
1967 Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues* 23(2). 29–38.

Gafaranga, Joseph
2000 Medium repair vs. other-language repair: Telling the medium of a bilingual conversation. *International Journal of Bilingualism* 4(3). 327–350.

Gafaranga, Joseph
2017 *Bilingualism as interactional practices*. Edinburgh: Edinburgh University Press.

García, Ofelia
2009 *Bilingual education in the 21st century: A global perspective*. Oxford: Wiley-Blackwell.

Gassner, Denise
2012 Vague language that is rarely vague: A case study of "thing" in L1 and L2 discourse. *International Review of Pragmatics* 4. 3–28.

Goffman, Erving
1959 *The presentation of self in everyday life*. New York: Anchor Books.
Goffman, Erving
1981 *Forms of talk.* Philadelphia: University of Pennsylvania Press.
Gumperz, John Joseph
1982 *Discourse strategies*. Cambridge: Cambridge University Press.
Haugen, Einar Ingvald
1953 *The Norwegian language in America*. Philadelphia: University of Pennsylvania Press.
Heller, Monica
2002 *Éléments d'une sociolinguistique critique*. Paris: Didier.
Hlavac, Jim
2014 Participation roles of a language broker and the discourse of brokering: An analysis of English-Macedonian interactions. *Journal of Pragmatics* 70. 52–67.
Horn, Laurence & Istvan Kecskes
2013 Pragmatics, discourse and cognition. In Stephen Anderson, Jacques Moeschler & Fabienne Reboul (eds.), *The language-cognition interface. Proceedings of the 19th International Congress of Linguists*, 353–373. Geneva: Droz.
Hymes, Dell
1972 Models of interaction of language and social life. In John J. Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics: The ethnography of communication*, 35–71. New York: Holt, Rinehart and Winston Inc.
Ifukor, Presley
2011 Linguistic marketing in "…a marketplace of ideas". *Pragmatics and Society* 2(1). 109–147.
Izadi, Dariush
2017 Semiotic resources and mediational tools in Merrylands, Sydney, Australia: The case of Persian and Afghan shops. *Social Semiotics* 7(4). 595–512.
Jaspers, Jürgen & Lian Malai Madsen
2016 Sociolinguistics in a languagised world: Introduction. *Applied Linguistics Review* 7(3). 235–258.
Jørgensen, J. Normann
2008 Polylingual languaging around and among children and adolescents. *International Journal of Multilingualism* 5(3). 161–176.
Karrebæk, Martha Sif
2017 Thai veggies and hair removal products: Space, objects and language in an urban green-grocery. *Social Semiotics* 7(4). 451–473.
Kress, Gunter
2009 *Multimodality: A social semiotic approach to contemporary communication*. London and New York: Routledge.
Kress, Gunther & Theo van Leeuwen
2001 *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.
Kusters, Annelies
2017 "Our hands must be connected": Visible gestures, tactile gestures, and objects in interactions featuring a deafblind customer in Mumbai. *Social Semiotics* 7(4). 394–410.
Lado, Robert
1957 *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press: Ann Arbor.
Li Wei
2011 Moment Analysis and translanguaging space: Discursive construction of identities by multilingual Chinese youth in Britain. *Journal of Pragmatics* 43(5). 1222–1235.

Li Wei
2018 Translanguaging as a practical theory of language. *Applied Linguistics* 39(1). 9–30.
Lo, Adrianne & Joseph Park
2017 Metapragmatics of mobility. *Language in Society* 46(1). 1–4.
Lorenzo-Dus, Nuria & Patricia Bou-Franch
2013 A cross-cultural investigation of email communication in Peninsular Spanish and British English: The role of (in)formality and (in)directness. *Pragmatics and Society* 4(1). 1–25.
Lou, Jackie Jia
2017 Spaces of consumption and senses of place: A geosemiotic analysis of three markets in Hong Kong. *Social Semiotics* 7(4). 513–531.
Lüdi, Georges & Bernard Py
2009 To be or not to be … a plurilingual speaker. *International Journal of Multilingualism* 6(2). 154–167.
Mondada, Lorenza (ed.)
2014 A body of resources: CA studies of social conduct. [Special issue]. *Journal of Pragmatics* 65.
Morris, Charles
1938 *Foundations of the theory of signs. International encyclopaedia of unified science*, vol. 1, no. 2. Chicago: University of Chicago Press.
Moya Guijarro, Arsenio Jesús
2011 Engaging readers through language and pictures: A case study. *Journal of Pragmatics* 43(12). 2982–2991.
Mugford, Gerrard
2013 Foreign-language users confronting anti-normative politeness in a Mexican university. *Intercultural Pragmatics* 10(1). 101–130.
Myers-Scotton, Carol
1993 *Social motivations for code-switching*. Oxford: Oxford University Press.
Okada, Yusuke
2015 Building rapport through sequentially linked joke-serious responses in second language job interviews. *Pragmatics and Society* 6(4). 593–614.
Otheguy, Ricardo, Ofelia García & Wallis Reid
2015 Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review* 6(3). 281–307.
Otsuji, Emi & Alistair Pennycook
2010 Metrolingualism: Fixity, fluidity and language in flux. *International Journal of Multilingualism* 7(3). 240–254.
Parvaresh, Vahid
2015 Vague language that is vague in both L1 and L2: A comment on Gassner (2012). *International Review of Pragmatics* 7(1). 129–143
Pennycook, Alastair & Emi Otsuji
2017 Fish, phone cards and semiotic assemblages in two Bangladeshi shops in Sydney and Tokyo. *Social Semiotics* 7(4). 434–450.
Poplack, Shana
1980 "Sometimes I'll start a sentence in Spanish y termino en español": toward a typology of code-switching. *Linguistics* 18(7–8). 581–618.
Rampton, Ben
2011 Style contrasts, migration and social class. *Journal of Pragmatics* 43. 1236–1250.
Reyes, Angela
2014 Linguistic anthropology in 2013: Super-new-big. *American Anthropologist* 116(2). 366–378.

Safont-Jordà, Maria-Pilar
2013 Early stages of trilingual pragmatic development: A longitudinal study of requests in Catalan, Spanish, and English. *Journal of Pragmatics* 59. 68–80.
Sambre, Paul & Kurt Feyaerts
2017 Embodied musical meaning-making, and multimodal viewpoints in a trumpet master class. *Journal of Pragmatics* 122. 10–23.
Schiffrin, Deborah
1996 Discourse analysis and pragmatics. In Hans Goebl, Peter H. Nelde, Zdenek Stary & Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact Linguistics / Linguistique de contact: Ein internationales Handbuch zeitgenössischer Forschung / An International Handbook of Contemporary Research / Manuel international des recherches contemporaines*, 138–144. Berlin: De Gruyter.
Selinker, Larry
1972 Interlanguage. *International Review of Applied Linguistics* 10(3). 209–241.
Tannen, Deborah
2005 *Conversational style: Analyzing talk among friends*. Oxford: Oxford University Press.
Ticca, Anna Claudia & Véronique Traverso
2017 Participation in bilingual interactions: Translating, interpreting and mediating documents in a French social centre. *Journal of Pragmatics* 107. 129–146.
Tseng, Ming-Yu
2016 Towards a pragmatic analysis of product discourse: Creative force and metapragmatic performance. *Pragmatics and Society* 7(1). 105–140.
Verschueren, Jef
1987 Pragmatics as theory of linguistic adaptation (IPRA working document 1). Antwerp: International Pragmatics Association.
Vertovec, Steven
2007 Superdiversity and its implications. *Ethnic and Racial Studies* 30(6). 1024–1054.
Weinreich, Uriel
1953 *Languages in contact: Findings and problems.* The Hague: Mouton.
Williams, Quentin
2017 Bark, smoke, and pray: Multilingual Rastafarian-herb sellers in a busy subway. *Social Semiotics* 7(4). 474–494.
Zhu Hua, Li Wei & Agnieszka Lyons
2017 Polish shop(ping) as translanguaging space. *Social Semiotics* 7(4). 411–433.
Zufferey, Sandrine, Willem M. Mak & Ted J. M. Sanders
2015 A cross-linguistic perspective on the acquisition of causal connectives and relations. *International Review of Pragmatics* 7. 22–39.

Emilee Moore, Barcelona (Spain)
Marco Santello, Leeds (UK)

13. Borrowing

1. Definitions
2. Types of borrowing
3. Motivations for borrowing
4. Structural integration
5. Constraints on borrowing
6. References

1. Definitions

The term 'borrowing' has been widely used in linguistic literature since the works of Haugen (1950) and Weinreich (1953) to refer to the adoption of a structure from one language within the framework of another. The language that is the source of the structure is usually referred to as the 'donor', while the language that has adopted it is referred to as the 'recipient'.

Borrowing is initiated when speakers of the recipient language come into contact with speakers of the donor language. Such contact can be superficial and limited to just a small group of users who are the importers of new terms, as in the case of English words like *tomahawk* or *sugar*, or it can be the property of an entire bilingual community. When users are bilingual and fluent in both the donor and recipient languages, the challenge is to identify lasting structural changes and to distinguish them from ad hoc language mixing or code-switching.

Poplack, Sankoff, and Miller (1988) argue that borrowing and code-switching are distinct phenomena. They defined 'borrowings' as structures that are formally integrated and show higher corpus frequency. These criteria are necessarily case-specific, since languages in contact may show structural similarities (such as phonological convergence) that can make it difficult to determine the degree of integration, and since setting a threshold for frequency is inevitably random to some extent. By contrast, Myers-Scotton (1993) treats all cases of language mixing as part of the same phenomenon, governed invariably by a hierarchical arrangement between a Matrix and an Embedded language.

A further dilemma concerns the status of Mixed Languages, where substantial structural components of two language systems have intertwined into one. Thomason and Kaufman (1988) characterize Mixed Languages as cases of 'broken transmission', arguing that there is no system continuity between the two source languages and the mixed outcome (see also Bakker and Matras 2003). This means that it is not possible to make a distinction between a donor and a recipient language. However, in later work, Thomason (2015: 40–41) and Thomason and Everett (2001) cite cases of Mixed Languages as examples of borrowing (of inflectional morphology, and of pronouns, respectively), thus assuming that we can, in fact, identify the donor and the recipient (see also Thomason and Kaufman's [1988: 103] characterization of Angloromani as a case of gradual, wholesale borrowing of English grammar into Romani, contrasting with the account offered in Matras [2010] of a historical shift from Romani to English, with selective retention of Romani lexicon as a stylistic device). This is of course crucial to our definition of 'borrowing': if Mixed Languages deserve their title as a category of cases that are fundamentally distinct from incremental, contact-induced language change, where items from a donor system are integrated into a recipient system, then they must remain outside the scope of a cross-linguistic corpus of borrowing examples, with implications for our

https://doi.org/10.1515/9783110435351-013

universal evaluation of borrowing, and the constraints that impact on it. I shall return to this point below.

Borrowing is most difficult to discern from code-switching in cases where active bilinguals alternate among meaning-equivalent structures in two (or more) languages. Conversely, the most unambiguous form of borrowing is the use of structures derived from another source language by an entire speech community, following a gradual process of propagation. The distinction between borrowing and code-switching is therefore best captured as a continuum: the more obvious case of borrowing is the default use of a structure for its inherent meaning (rather than for conversational effect), on a regular basis (rather than ad hoc), and by speakers who do not have active command of the language that is the historical-etymological source of the structure in question (Matras 2009: 110–114, 146–149).

2. Types of borrowing

A variety of concepts have been used to capture different forms of cross-language replication of structures, and I will mention just some of these. A distinction can be made between replication of linguistic material in the form of a phonological shape with associated meaning, and that in the form of a structural configuration in which the position and order of elements, and the relations among them, convey a certain meaning. Haugen (1950) and Weinreich (1953) use the term 'borrowing' to refer to the former, while the latter was labelled 'calque' or 'convergence'. The difference can be illustrated by the example of Sinti Romani (spoken in Germany and neighboring regions), where *aktionsart* modifiers to the verb replicate a German model. The phrase *me kerau pre*, literally 'I make up', replicates German *ich mache auf* meaning 'I open'. It draws on the inflected Romani verb *ker* 'to do' and the preposition *pre* 'up', and replicates the meaningful alignment of the corresponding structures in the German construction *mache* 'I do' and *auf* 'up, on', exploiting the polysemy of German *auf*, which is also an adjective meaning 'open'. By contrast, German *ich gehe hin* 'I am going (to an aforementioned destination)' is replicated in Sinti Romani as *me džau hin*. Here, the Romani verb *džau* 'I go' is accompanied by the German borrowing *hin* 'towards'.

Strictly configurational features such as word order are by necessity candidates only for calquing, convergence, or 'pattern' replication. But as the Sinti Romani example demonstrates, in lexicon and morphology both borrowing (replication of linguistic 'matter') and calquing (replication of 'pattern') are in principle possible. There are, however, some tendencies that point to a preference among certain categories, though these are not strict constraints: case and agreement morphemes, for example, are rarely borrowed as matter-forms, but the meaning and distribution of case markers can be the subject of convergence among languages, with one replicating the system of another by drawing on inherited forms (Tenser 2016). Definite articles, too, are often subject to convergence. Romani replicates Byzantine Greek preposed articles, deriving them from its inherited inventory of demonstrative pronouns (Matras 2002: 96), and the Sorani Kurdish definite article *eke* is regarded as a replica of the article of contiguous Semitic languages (notably Syriac or Aramaic), while the very same suffixed article is then borrowed from Kurdish in its matter-form into modern Aramaic (Khan 2007).

Another distinction relates to the diffusion of a borrowing within the receiving system. English replicates the Latin inflectional distinction between singular and plural in *fungus* versus *fungi*, but it remains unproductive and limited to a small number of Latin nouns, diffusing neither 'backwards' to inherited English words, nor 'forwards' to other loanwords or new vocabulary. By contrast, Yiddish borrows the Hebrew plural marker with Hebrew loans such as *xavéjrim* 'friends', which diffuses forwards to some European loans such as *doktójrim* 'doctors'. Kurdish borrows the pharyngeal phoneme /ḥ/ in Arabic loans such as *ḥeywan* 'animal', which diffuses backwards into inherited Kurdish words like *ḥeft* 'seven' (< *heft*), substituting for an original glottal.

Borrowing may involve the adoption or replication of a single feature, or of a set of features, with potential implications for the typological arrangement of the receiving system. Phonetic and phonological features can accompany individual loanwords, but may also, as in the above example, become part of the inherited inventory, with implications for the recipient language's sound system as a whole. In Romani, Modern Greek affricates /dz/ and /ts/ are largely limited to borrowed words, while in Kurdish, as mentioned above, pharyngeal sounds, which originally only accompanied Arabic loans, can appear in inherited words. Borrowing can also entail a loss of features, in which case the notion of replication can only be applied to the system as a whole and not to individual structures: Russian Romani has lost the historical Romani definite article by replicating the Russian system with no overt marking of definiteness.

Replication of configurational patterns can also have implications for the typology of the system as a whole. Rumelian Turkish shows finite adverbial subordinations, and finite complement and relative clauses, as well as a formal distinction between factual and non-factual complements, consistently matching the structures of neighboring languages in the Balkans (Bulgarian-Macedonian, Albanian, Greek; Matras and Tufan 2007). Here one can speak of overall convergence among the languages, as it is difficult to identify precisely which contact language was the source of a particular construction, or indeed whether the similarities might be due to a generalization of finite constructions that had been inherited from Oghuz Turkish and had existed alongside nominal-participial constructions (these became the norm in Ottoman Turkish). While the classic concept of 'borrowing' pertains, as noted above, to an identifiable structure with identifiable etymology, diachronic analyses often need to adopt a broader interpretation that would include the possibility of a contact-induced preference to select a particular structure among inherited options, or even just to accelerate internal change and alter the distribution of a construction. Heine and Kuteva (2005) therefore speak of contact-induced grammaticalization as a driver of morpho-syntactic change in contact situations. They propose that general notions of grammaticalization theory such as semantic extension, expansion of distribution, and increased frequency, can account for and help frame processes of convergence. The issues of 'belonging' to an original donor system and 'employment' in the receiving system are dealt with by identifying one of the language-specific constructions as a 'model', and the other as a 'replica'.

Another way of approaching the question of borrowing in cases of morpho-syntactic convergence is the notion of 'pivot-matching' (Matras 2009: 240–242). This is compatible in principle both with the idea of contact-induced grammaticalization, and with that of a replication of constructions (a topic dealt with in several of the contributions to Wiemer, Wälchli, and Hansen [2012]), but is not constrained by the directionality condition of the grammaticalization model. This allows us to account both for the loss of

categories as a result of contact and for the acquisition of new ones, and it can offer a local account of the actual construction components that are affected by cross-language replication. Consider the following examples from Khuzistani Arabic (Matras and Shabibi 2007):

(1) Khuzistani Arabic:
a. *walad č-čibīr*
boy DEF-big
'the big boy'

b. *walad l-modīr*
boy DEF-director
'the director's son'

Attributive constructions involving adjectives (1a) and modifying nouns (1b) follow essentially the same pattern, with the historical Arabic definite article *(a)l* (which assimilates to dental/palatal consonants, resulting in consonant germination in [1a]) functioning as a nominal attributive marker taking on a proclitic position with the attribute. This replicates the Persian system, with which Khuzistani Arabic is in contact, where the nominal attributive ending *e* serves the same function in the same position in the noun phrase:

(2) Persian:
a. *pesar-e bozorg*
boy-ATT big
'the big boy'

b. *pesar-e modīr*
boy-ATT director
'the director's son'

By contrast, Classical Arabic, Modern Standard Arabic, and all other contemporary Arabic dialects distinguish between adjectival attribution, where noun and adjective agree in definiteness, and nominal attribution, where the definite article links the NP with its possessor-modifier:

(3) Arabic (general):
a. *l-walad l-kabīr*
DEF-boy DEF-big
'the big boy'

b. *walad l-mudīr*
boy DEF-director
'the director's son'

What is 'borrowed' from Persian into Khuzistani Arabic is the absence of a contrast between the two constructions. At the core of the process lies an interpretation of the pivotal feature of the Arabic attributive constructions − the mediating definite article in

the position between head and modifier − as matching, in scope and function, the pivotal feature of the corresponding Persian construction − the attributive (ezafe) marker *e*.

The notion of 'heavy borrowing', introduced by Thomason and Kaufman (1988) and often cited in specialized literature, presupposes either an agreed-upon threshold for borrowing (which is not available), or the borrowing of categories that are less prone to cross-linguistic replication (which in turn would suggest a hierarchical understanding of borrowing, a point to which I return below). The following example from Palestinian Domari offers an uncontroversial example of heavy borrowing:

(4) Palestinian Domari:
kān ʕumrom yimkin sitte snīn sabʕa snīn
was.3SG.M age1SG maybe six years seven years
'I was maybe six or seven years old'.

Here, all items except for the 1SG possessive ending *om* in the word *ʕumrom* 'my age' are derived from and shared with the contemporary contact language, Palestinian Arabic (Matras 2012). This includes the nouns 'age' and 'year' (along with the latter's plural marking), the modal particle 'maybe', the numerals, and, most exceptionally for borrowing, the auxiliary verb 'to be' along with its Arabic inflection. If this Domari utterance might be considered a yardstick for heavy borrowing, then it would be the borrowing of particular categories, most notably the numerals and the inflected auxiliary, rather than the random density of Arabic-derived structures in one particular utterance, that would constitute a measure of the degree of borrowing and its impact on the receiving system, and would thus in this case qualify as 'heavy'. But the overall density of Arabic loans in this utterance does provide an illustration of just how far-reaching the structural implications of borrowing can be.

3. Motivations for borrowing

Two frequently cited motivations for borrowing are 'gaps' in the recipient system, and the 'prestige' of the donor system. Gaps of course explain the borrowing of terms for new artefacts, products, and cultural innovations (such as English *banana*, *sushi*, or *parliament*), including social and ideological concepts (such as English *majesty* or *redemption*). Comparative data on loanwords in a sample of languages (Haspelmath and Tadmor 2009; see also World Loanword Database: http://wold.clld.org/) shows some clear hierarchical trends, with semantic categories, such as modern world artefacts, leading in proneness to borrowing, followed by religion and belief, clothing and grooming, household items, and the law. Prestige, by contrast, is a vague notion. It tends to refer to the dominant status of a language within a particular domain of social interaction. Thus, domain-specific borrowings, such as English *beef*, *mouton*, *poultry*, and *pork*, which complement inherited terms, like *ox*, *sheep*, *chicken*, and *pig*, might be explained as reflecting the role of French as a preferred language of the medieval English elite and a source of imitation in connection with culinary customs. Similarly, the English slang expressions *pal* 'friend', *kushty* 'good', *chavy* 'guy', and *minge* 'vagina' from Romani

reflect the association of Romani in eighteenth and nineteenth century England with defiance of the authority of the establishment.

The motivation behind such borrowings is thus invariably connected to the specialized roles of languages in different social interaction domains. Thus, Khuzistani Arabic has *bāṣ* 'bus', a pan-Arabic loan from (colonial) English, in the singular, but *otobushā* 'busses', from Persian (which in turn borrows the singular form from French), in the plural, reflecting the contemporary dominance of Persian in the domain of formal-institutional communication, where busses are more likely to be referred to in the plural. Hebrew, by contrast, has the European loanword *oto* 'car' in the singular but the neologism *mexoniyot* 'cars' in the plural, a reflection of the impact of language-engineering and revivalist ideologies on formal-institutional communication domains.

Neither gaps nor prestige, however, explain borrowings that replace inherited items, or the apparent hierarchical nature of borrowing in some cases. For example, borrowed terms are more likely to target more remote kin (English *uncle*, *aunt*, *grandparents*, *niece*, *nephew*, from French; Maltese *nannu* 'grandfather, *ziju* 'uncle', *kuġin* 'cousin', *neputi* 'nephew' from Italian/Sicilian) than proximate kin words for parents, siblings, and direct offspring; and languages that borrow numerals, such as Swahili and Palestinian Domari (both from Arabic) and Romani (from Greek) are more likely to borrow numerals above '5' than under. There are thus, in addition to the effect of diglossia, also cognitive factors in operation, whereby everyday concepts that are simple, close, intimate, and more frequently used tend to resist borrowing, whereas their paradigm counterparts that indicate complexity, remoteness, formality, and tightly-regulated routines are more borrowing-prone. From this one might draw the conclusion that borrowing can serve, in its incipient stage, as a meaningful way of marking asymmetry between paradigm values (Elšík and Matras 2006: 385).

The question arises in particular in relation to the borrowing of functionally equivalent (or near-equivalent) structures in grammar. Poplack (1980) argues that the frequent insertion of English discourse markers into the Spanish discourse of semi-bilingual Puerto Ricans in New York City serves to flag the prestige of being bilingual. But for fluent bilinguals, the tendency to replace such items with those derived from the language that dominates in formal interaction contexts (the prestige language) can be seen as an effort to reduce cognitive processing load; this is evidenced both through bilingual speech production errors that are in effect 'counter-prestigious', and through the hierarchical nature of historical borrowing that shows items that challenge presuppositions at the top of the susceptibility scale (Matras 1998, 2000). An attempt to provide a more holistic approach is Myers-Scotton's (2002) '4M-model', which postulates a stratification of morphemes in the speech production process and links it to their likelihood to be affected by language mixing. One can certainly assume that factors such as 'executive control' (Green 1998) over the selection of items in the repertoire will have a differentiated impact on grammatical categories, in effect making it more or less of a challenge for bilingual speakers to maintain system separation around particular categories, thereby influencing the likelihood of longer-term generalization across the repertoire of just one construction, and the subsequent propagation of that construction throughout the speech community, leading to long-term borrowing (Matras 1998).

4. Structural integration

According to some definitions (e.g. Poplack, Sankoff, and Miller 1988), borrowings are always structurally integrated into the recipient system, meaning that they are phonologically and morphologically adapted, or that they take on the 'system morphemes' and linear order rules of the receiving, or 'matrix', language (Myers-Scotton 1993, 2002). Yet the reality is more complex. Phonological adaptation may be characteristic, as in Japanese *pikunikku* 'picnic', but the retention of the Arabic pharyngeal in Kurdish *ḥeywan* 'animal' (an Arabic loanword) shows that borrowed words can also be a conduit for borrowed phones and phonemes. Morpheme boundaries in the donor language are not always analyzable in the process of morphological integration, as seen in Spanish *aceite* 'oil' from Arabic *az-zēt* '(the) oil', where the original definite article is interpreted as part of the noun's generic meaning, or conversely, Swahili *kitabu* 'book', *vitabu* 'books', where the initial syllable in the Arabic word *kitāb* is reinterpreted as a Swahili nominal classifier *ki*. Some languages show morphological compartmentalization, with borrowed lexical items taking borrowed morphology. In Romani, Greek-derived verbs, like *jir(iz)* 'to return' and *xa(s)* 'to lose', retain their Greek-derived aspectual endings, which are also generalized to all subsequent (post-Byzantine) loan verbs, as in Turkish-derived *anladiz* 'understand' and Slavic-derived *končis* 'to complete' (Matras 2002: 128), while in some Romani dialects, Turkish loan verbs retain their complete Turkish inflection, as in Kalburdžu Romani (Sindel, eastern Bulgaria) *mangav* 'I want' (inherited Indic), but *konušuim* 'I speak' (borrowed, Turkish). In Palestinian Domari, modal and auxiliary verbs retain their complete Arabic inflection, as in *ṣār* 'he began', *ṣirt* 'I began' (borrowed, Arabic), but *šarda* 'he said', *šardom* 'I said' (inherited, Indic).

Borrowed verbs are an interesting case. Structural integration patterns range from no particular marking, as in English *demand*, to historical changes in adaptation patterns, as in German *telefonieren* 'to telephone', *protestieren* 'to protest', borrowed from French in the nineteenth century, but *parken* 'to park', *downloaden* 'to download', borrowed from English since the late twentieth century. Many languages use particular devices such as: a distinct verb inflection class, as in the Romani examples cited above; an inflection class that conveys causativity, intensitivity, or iterativity, as in Hebrew *tilfen* 'to telephone' and *kitleg* 'to catalogue', akin to *kirev* 'to draw closer', *šiber* 'to break into pieces', and *šitek* 'to paralyse'; or explicit verbalization by means of adding a light verb, as in Turkish *teşekkür etmek* 'to thank' (from Arabic/Persian *tašakkur* + *etmek* 'to do'). The typology of verb borrowing thus shows a range of options that indicate a continuum in the extent to which a lexical item's 'verbness' is taken for granted based on its meaning, or requires explicit flagging through either inflection or composition. This in turn can be taken to reflect the cognitive complexity of verbs as carriers of meaning, deictic and presupposition features (tense, modality), internal event structures (aspect), and argument structures (Matras 2009: 175; Wohlgemuth 2009).

5. Constraints on borrowing

A common view in contact linguistics is that borrowing is conditioned by the structural congruency of the specific languages in contact, and by the sociolinguistic relations

among them. Yet not every instance of borrowing is predictable, even in cases where we have identical prerequisites in these two areas. For example, Jordanian Domari borrows the Arabic (feminine) plural ending *āt* and uses it with inherited Indic nouns (backwards-diffusion), irrespective of grammatical gender, as in *putrāt* 'children', while Palestinian Domari only shows the ending with (feminine) Arabic loans and retains inherited plural formation with inherited nouns, as in *putre* 'children'. Both varieties have an identical history of contact with surrounding languages, under the same sociolinguistic conditions, and are structurally the same language with only very minor dialect differences.

In searching for predictions and identifying constraints on borrowing, we can be guided instead by two theoretical considerations: the postulation of meaningful implicational hierarchies of borrowing, and the rareness of borrowing around certain functional or structural categories. These are perhaps the most fascinating and rewarding, but also the most controversial, and empirically most challenging, areas of research on borrowing.

Notional implicational hierarchies were proposed by Moravcsik (1978) in relation to specific pairs of typological categories, such as lexical > non-lexical, nouns > non-nouns, free morphemes > bound morphemes, and derivation > inflection, rendering the overall theme of semantic transparency a facilitator for borrowing (or non-transparency a restriction); and by Thomason and Kaufman (1988) in the form of a gradient of likelihood of borrowing, in which, however, some categories such as 'function words' or 'typological features' remain rather vague and undifferentiated, while the overall theme merely suggests that prolonged and intensive contact between languages is likely to yield more extensive borrowing. More specific results have since been obtained through targeted, cross-linguistic sampling (Elšík and Matras 2006; Matras 1998; Matras and Sakel 2007; Stolz and Stolz 1997). This work points to a connection between susceptibility to borrowing, and the truth- or presupposition value assigned by a category to propositional content. A good example is the borrowability hierarchy of connectives 'but > or > and' (Matras 1998) (where 'greater than' indicates greater likelihood of borrowing, and the order is implicational), a hierarchy that has been widely attested across different samples. Here, contrast − the unexpected that is beyond the speaker's control and therefore associated with interactional tension, as it puts the speaker's assertive authority potentially in jeopardy − is more prone to borrowing than addition, which conveys a continuous and expected inferential chain. Similar hierarchies have been identified for other functions, such as indefinites, comparatives, tense and modality, and more (Matras 2007, 2009: 218). Such hierarchies are difficult to explain, especially from a cross-linguistic perspective, with reference to either formal-structural features, or merely to sociolinguistic conditions and prestige. Rather, they capture an interactional dimension, whereby those categories that clash with, or cannot be firmly derived from, presuppositions are more prone to borrowing. This suggests that the speaker's management and control of the interaction is linked to the speaker's executive control of repertoire components, and that in turn the diachronic process of borrowing is linked to the processing of language in conversation in bilingual settings.

Identifying cross-linguistic borrowing universals is an empirical challenge, not least because of the need to have information about the etymology of structures and items, as well as about the history of contacts, information that is not available for most languages of the world. Even where documentation of contact phenomena exists, it is often incomplete, and selective reading of isolated observations can easily lead to misinterpretations. For example, the mention, in a first-hand source, that a Turkic language borrowed the

Iranian comparative suffix *tar*, does not imply that it is only the comparative, and not the superlative, that is borrowed (*pace* Seifart 2017), and in fact, Iranian *tar* serves as both comparative and superlative. When such errors of interpretation infiltrate a statistical analysis that is based on a rather small sample, then they risk invalidating the conclusions.

It is also empirically difficult (or even impossible) to rule out that some items or categories can be borrowed, or to claim with certainty that they cannot. Yet the extensive documentation that is available on borrowing does show that there are some significant differences in frequency among borrowed categories, both in absolute terms, and in relation to the borrowing of other categories in the same language contact setting. Attestations of borrowed pronominal forms, for example, or of bound person morphemes, are relatively rare. Thomason and Everett (2001) try to challenge the assumption that pronouns are rarely borrowed. However, as examples they cite Mixed Languages. This undermines the very definition of Mixed Languages as examples of 'broken transmission' (see above). Following Wallace (1983), they also cite cases from Southeast Asia. However, Wallace himself had identified these cases as functionally and typologically distinct from Indo-European pronouns, in that they essentially convey lexical meaning, and are thus more akin to expressions such as 'Your Majesty' or 'Your Honor', rather than constituting deictic or anaphoric elements. Indeed, Thomason (2015) mentions the ambiguity of Mixed Languages as examples of the borrowability of inflectional morphemes, for which attestation is equally rare.

The view put forward by Thomason (2015), and recently reiterated by Seifart (2017), that there are no strict constraints on borrowing, might be seen as representing a rejection of an epistemology according to which the purpose of investigating borrowing, and indeed of any scientific investigation, is to venture beyond the mere listing of facts and to attempt instead to offer an explanatory account of those facts. For its part, the search for constraints on borrowing, and for the factors that facilitate borrowing, rests on an epistemology which identifies overwhelming trends as worthy of attention (even if isolated exceptions exist), which purports to be able to derive explanatory models from trends, and which indeed sees the purpose of the cross-linguistic examination of borrowing as an ambition to be able to formulate an explanatory model. It operates on the assumption that borrowing is not random, but that it is conditioned in some way by human behavior, and that it reflects the dimensions of human communication and the processing of information and knowledge that are encoded in the structural categories of language. In this respect, the study of borrowing hierarchies has the potential to provide an important window into a better understanding of the layered structure of the language faculty itself, and possibly even of key aspects of the evolution of human language.

6. References

Bakker, Peter & Yaron Matras (eds.)
2003 *The mixed language debate: Theoretical and empirical advances*. Berlin: Mouton de Gruyter.

Elšík, Viktor & Yaron Matras
2006 *Markedness and language change: The Romani sample*. Berlin: Mouton de Gruyter.

Green, David
1998 Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1(2). 67–81.
Haspelmath, Martin & Uri Tadmor (eds.)
2009 *Loanwords in the world's languages: A comparative handbook*. Berlin: Mouton de Gruyter.
Haugen, Einar
1950 The analysis of linguistic borrowing. *Language* 26. 210–231.
Heine, Bernd & Tanya Kuteva
2005 *Language contact and grammatical change*. Cambridge: Cambridge University Press.
Khan, Geoffery
2007 Grammatical borrowing in North-Eastern Neo-Aramaic. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 197–214. Berlin: Mouton de Gruyter.
Matras, Yaron
1998 Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36(2). 281–331.
Matras, Yaron
2000 Fusion and the cognitive basis for bilingual discourse markers. *International Journal of Bilingualism* 4(4). 505–528.
Matras, Yaron
2002 *Romani: A linguistic introduction*. Cambridge: Cambridge University Press.
Matras, Yaron
2007 The borrowability of grammatical categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 31–74. Berlin: Mouton de Gruyter.
Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.
Matras, Yaron
2010 *Romani in Britain: The afterlife of a language*. Edinburgh: Edinburgh University Press.
Matras, Yaron
2012 *A grammar of Domari*. Berlin: De Gruyter Mouton.
Matras, Yaron & Jeanette Sakel (eds.)
2007 *Grammatical borrowing in cross-linguistic perspective*. Berlin: Mouton de Gruyter.
Matras, Yaron & Maryam Shabibi
2007 Grammatical borrowing in Khuzistani Arabic. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 134–149. Berlin: Mouton de Gruyter.
Matras, Yaron & Şirin Tufan
2007 Grammatical borrowing in Macedonian Turkish. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 215–227. Berlin: Mouton de Gruyter.
Moravcsik, Edith
1978 Universals of language contact. In Joseph H. Greenberg (ed.), *Universals of human language*, 94–122. Stanford: Stanford University Press.
Myers-Scotton, Carol
1993 *Duelling languages: Grammatical structure in codeswitching*. Oxford: Oxford University Press.
Myers-Scotton, Carol
2002 *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.

Poplack, Shana
1980 Sometimes I'll start a sentence in Spanish y termino en español. *Linguistics* 18. 581–618.
Poplack, Shana, David Sankoff & Carol Miller
1988 The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26. 47–104.
Salmons, Joseph
1990 Bilingual discourse marking: Code switching, borrowing, and convergence in some German-American dialects. *Linguistics* 28. 453–480.
Seifart, Frank
2017 Patterns of affix borrowing in a sample of 100 languages. *Journal of Historical Linguistics* 7(3). 389–431.
Stolz, Christel & Thomas Stolz
1997 Universelle Hispanismen? Von Manila über Lima bis Mexiko und zurück: Muster bei der Entlehnung spanischer Funktionswörter in die indigenen Sprachen Amerikas und Austronesiens. *Orbis* 39. 1–77.
Tenser, Anton
2016 Semantic map borrowing: Case representation in northeastern Romani dialects. *Journal of Language Contact* 9(2). 211–245.
Thomason, Sarah G.
2015 When is the diffusion of inflectional morphology not dispreferred? In Franceso Gardani, Peter Arkadiev & Nino Amiridze (eds.), *Borrowed morphology*, 27–46. Berlin: De Gruyter Mouton.
Thomason, Sarah G. & Daniel L. Everett
2001 Pronoun borrowing. In Charles Chang, Michael J. Houser, Yuni Kim, David Mortensen, Mischa Park-Doob & Maziar Toosarvandani (eds.), *Proceedings of the twenty-seventh annual meeting of the Berkeley Linguistics Society*, 301–315. Berkeley: Berkeley Linguistics Society.
Thomason, Sarah G. & Terrence Kaufman
1988 *Language contact, creolization and genetic linguistics.* Berkeley: University of California Press.
Wallace, Stephen
1983 Pronouns in contact. In Frederick B. Agard, Gerald Kelley, Adam Makkai & Valerie Becker Makkai (eds.), *Essays in honor of Charles F. Hockett*, 573–589. Berlin: Mouton de Gruyter.
Weinreich, Uriel
1953 *Languages in contact*. The Hague: Mouton.
Wiemer, Björn, Bernhard Wälchli & Björn Hansen (eds.)
2012 *Grammatical replication and borrowablity in language contact*. Berlin: De Gruyter Mouton.
Wohlgemuth, Jan
2009 *A typology of verbal borrowings*. Berlin: Mouton de Gruyter.

Yaron Matras, Manchester (UK)

14. Code-switching

1. Historical overview and terminology
2. Linguistic perspectives on code-switching
3. Sociolinguistic perspectives on code-switching
4. Acquisition and competence perspectives on code-switching
5. Conclusion
6. References

1. Historical overview and terminology

The alternation between languages in speech started receiving specific attention in the post-war years with Weinreich (1953). 'Code-switching' – as Vogt first called it in 1954 – developed into a field cutting across structural linguistics, sociolinguistics, and psycholinguistics, with boundaries between code-switching proper, and language variation in general, becoming porous. I refer to 'code-switching' as the overt alternation between languages or varieties. This sets code-switching against the diachronic outcomes of contact known as 'borrowings', 'transfers', and 'focusing', as well as against the concepts of 'cross-linguistic influence', 'transfer', and 'interference' used in language acquisition research.

The theorization of code-switching (CS) per se began with Blom and Gumperz (1972) and Gumperz (1982) establishing the notion that CS performs specific social functions across settings. At a later stage, scholars began to search for universal grammatical constraints to CS, with Poplack (1980) and Sankoff and Poplack (1981), who, based on surface structures, proposed the Free Morpheme Constraint and Equivalence Constraint. Subsequent constraints were formulated based on deep structures, such as the Government Constraint (Di Sciullo, Muysken, and Singh 1986), Functional Head Constraint (Belazi, Rubin, and Toribio 1994), and the constraints in Myers-Scotton's (1993a) Matrix Language Frame model. In contrast, MacSwan (1999, 2005) proposed that nothing constrains CS, apart from properties of mixed grammars, while Muysken (2000) accounts for grammatical regularities in CS through both potential grammatical universalia, and factors linked to prestige and competence levels.

Social context and social structure are evaluated differently by sociolinguists as explanations for CS. For the Communication Accommodation Theory (Gallois, Ogay, and Giles 2005) and the Markedness Model (Myers-Scotton 1993b), CS, and the social meanings that it conveys, is determined by norms of interaction. By contrast, Gumperz (1982) favors ethnographic views of multilingual behaviors as forming situational identity negotiation strategies through which norms can be confirmed or undone. Conversation Analysis (CA), finally, views CS as a marker of conversational structure rather than of social structure (Auer 1998).

Many approaches to CS assume that languages form bounded units that overlap, subject to conditions, such as that only one language sets the grammatical frame for bilingual clauses (Myers-Scotton 1993a). The emphasis placed on interlingual boundaries often also facilitates a distinction between CS as an ad hoc phenomenon, and long-term, conventionalized outcomes of CS in the form of established borrowings – which involve isolated lexical elements – and transfers – which may involve borrowings and contact-induced grammatical changes (Clyne 2003; Thomason 2001; see further Treffers-Daller [2009] and Chapter 2 in this volume). The fact that intensive borrowing and

https://doi.org/10.1515/9783110435351-014

transfer can give rise to 'mixed languages' has helped qualify claims that languages remain separate in interaction, and has ignited interest in 'ungrammatical' CS, described as, among other things, 'congruent lexicalization' (Muysken 2000). In the ethnographic paradigm, contact phenomena have been placed in the perspective of (stylistic) repertoires, providing the bedrock for new terminological developments, including the terms 'polylanguaging' and 'translanguaging', which now compete with the term CS (Blommaert 2010).

This interest in the psychological dimensions of bilingualism gives CS a place in the study of (second) language acquisition (SLA), and of broader neuro-cognitive processes. Research on both early language and L2 acquisition characteristically refers to 'transfers' and 'interferences', which may include CS (see Chapter 15). Psycholinguistic models of bilingual production have resorted to the concept of 'language activation' as a generic explanation for CS (Clyne 2003; de Bot 1992). Both paradigms use CS to produce insights into the cognitive separation between the languages of bilinguals.

This chapter summarizes three approaches to CS: Section 2 treats the linguistic perspective, Section 3 the sociolinguistic perspective, and Section 4 the psychological perspective.

2. Linguistic perspectives on code-switching

2.1. Universal constraints on code-switching

From the 1980s onwards, studies of CS became ever more associated with the search for universal grammatical rules in its forms, as a growing amount of CS data suggested regular structural patterns in CS across language contact situations. The earliest attempt at formulating universal grammatical constraints of CS was made by Poplack (1980), who subsumed structural regularities in surface patterns of Spanish-English CS among New York Puerto Ricans into two grammatical constraints, namely: the Free Morpheme Constraint, which prohibits switching between free and bound morphemes (such as in the form **eat-iendo*, where the Spanish bound morpheme *-iendo* is 'ungrammatically' affixed to the English stem *eat*), and the Equivalence Constraint, which prohibits switching where the two participating languages display no linear equivalence.

Other contact situations provided counterevidence to the universal character of Poplack's constraints (Clyne 1987; Nortier 1990). While Poplack justified her constraints based on statistical distributions, subsequent constraints took Universal Grammar principles as a point of departure. The Government Constraint (Di Sciullo, Muysken, and Singh 1986) holds that the language of the head remains that of its maximal projection (e.g. a verb must be in the language of its complement). A subsequent influential, government-based constraint was the more specifically phrased Functional Head Constraint (Belazi, Rubin, and Toribio 1994), which prohibits CS between functional heads (as opposed to lexical heads) and their complements. A point of criticism recurrently raised over such constraints, as well as over Poplack's Equivalence and Free Morpheme constraints, is that they leave much room for (or in the latter case, openly posit) mechanisms specific to CS, that is, mechanisms operating independently of the grammars of the participating languages.

Avoiding the need for a 'Third Grammar' (Pfaff 1979) has been a consistent motivation behind the formulation of models of CS within the Universal Grammar paradigm, as exemplified by Woolford (1983), Mahootian (1993), and Belazi, Rubin, and Toribio (1994). A new scope for fulfilling that motivation came with the Minimalist Program (MP; Chomsky 1995). By positing that grammatical rules are mediated by lexical items, the MP facilitates a view of CS, most prominently articulated by MacSwan (1999), as subject to no syntactic constraints other than those inferable from the grammars of the participating languages. The constraints on CS that MacSwan (2014) proposes concern, among other things, CS in word-internal contexts, CS between lexical items which cannot check one another's grammatical features (e.g. in the Spanish-English example **the casa*, where the English determiner cannot reflect the gender feature carried by Spanish *casa*), and the syntactic position of switches (e.g. the language of the verb determines the position of the subject).

Whereas MacSwan posits that lexical choices are made before the syntactic structure of the sentence is laid out, the Matrix Language Frame model (MLF; Myers-Scotton 1993a) assumes that lexical choices are constrained by a predefined morphosyntactic frame, which only one language (the Matrix Language [ML]) supplies in bilingual speech. The MLF differs from other grammatical approaches to CS in that it refers to 'language activation' (see Section 4) to justify the asymmetric distinction it makes between the ML (i.e. the more activated language), and the Embedded Language (EL), which is found in the form of 'language islands' in the ML. MLF also posits a hierarchy of insertable elements from Language Y into Language X, articulated around a distinction between content morphemes and system morphemes, the latter being less easily insertable from Language Y into Language X outside of language islands (see MacSwan [2005] for a critique).

Myers-Scotton acknowledged that the MLF cannot predict all grammatical forms of CS. This provided a justification for the more descriptive approach proposed by Muysken (2000, 2013), who acknowledges four grammatical types of CS: 'insertional CS' involves other-language constituents embedded in one dominant morphosyntactic frame, 'alternational CS' occurs at boundaries between clausal units, 'congruent lexicalization' describes CS whereby the morphosyntactic frame cannot be unambiguously assigned to one specific language, and 'back-flagging', finally, refers to the combination of discourse markers from Language X with utterances in Language Y (Muysken 2013). Instead of phrasing grammatical constraints, Muysken observes trends: CS involving typologically close languages is more likely to produce insertional CS or congruent lexicalisation, while CS involving typologically distant languages is more likely to produce alternational CS.

2.2. Code-switching, borrowings, and transfer(ence)s

Determining what is (not) a switch in bilingual speech has been a consistent concern for structural approaches. Constraint-based models distinguish between CS and borrowing. A widespread early view (Poplack 1980) saw phonological integration as a feature of borrowings, and a condition for lexical items from a donor language to be fit into the morphological frame of the host language (thus escaping the Free Morpheme Constraint). Poplack and her associates subsequently observed that some infrequently occurring donor language lexical items do not necessarily exhibit phonological adjustment to the host

language, while behaving in accordance with its morphological rules; this led them to the concept of 'nonce borrowing', an intermediate category of donor language items located between code-switches and borrowings (Poplack and Meechan 1995; Poplack, Sankoff, and Miller 1988).

While accepting a disjunction between phonological and morphological integration, Myers-Scotton (1993a, 2002) contested the rationale behind nonce borrowings, as well as the relevance of an analytical distinction between borrowing and CS: both borrowings and single-item code-switches are similarly subject to ML morphological procedures, except in the case of single-item code-switches forming EL islands. In contrast, Muysken (2000) retains an analytical distinction between borrowing and CS, implying that the two may display differential levels of integration. One predictor of integration is whether the donor language item is inserted at a supra- or sublexical level. The other is the frequency-related criterion of 'listedness', i.e. the "degree to which a particular element or structure has become part of a memorized list that has gained acceptance within a particular speech community" (Muysken 2000: 71).

A theme in discussions of borrowing is the diachronic effect of CS on the host language's lexicon and structure (see Chapter 13). Van Coetsem (1988) and Thomason and Kaufman (1988) laid a foundation for this in historical linguistics. Another point of reference was provided by Muysken's (1981) 'Relexification Hypothesis' in relation to Media Lengua (an Andean language with a Spanish lexicon and a Quechua grammatical frame). The concept of 'mixed' languages (Thomason and Kaufman 1988; Bakker 1997) provided a background for specifically discussing the long-term grammatical effects of CS. Myers-Scotton (1993a) was first to explicitly place the issue within the perspective of CS via her 'Matrix Language Turnover Hypothesis'. Frequent insertional CS can lead to a composite ML straddling the distinctions between the initial ML and EL.

An important factor in ML turnovers, and change in general, is convergence between two interacting languages, often discussed as a phenomenon connected with language shift. Clyne (2003) describes convergence as a form of 'transference': the process of transferring elements and patterns from one language to another in a (generally shifting) bilingual community. Convergence, however, may occur without transference, as illustrated by 'attrition', in which a feature present in one language, but not in another, may disappear (also known as 'levelling'). A major question has been whether CS creates a conduit for convergence, or the other way around. It is now generally accepted that CS, although it is a potential conduit for convergence, may just as well not bear any relation to it (Clyne 2003; Myers-Scotton 2002; Thomason 2001).

Convergence has been described not only between languages, but also between closely related varieties, such as 'dialects', that is, varieties "used in a geographically limited part of a language area in which they are 'roofed' by a 'structurally similar standard variety" (Hinskens, Auer, and Kerswill 2005: 1). Studies of dialect contact often use variationist methodologies to monitor diachronic change, such as 'koineization', a process which involves 'dialect mixing', 'levelling', and 'simplification' (Trudgill 2006). The term 'focusing' was originally introduced by Le Page and Tabouret-Keller (1985) to capture the reduction of free variation between semantically equivalent linguistic features in inter-lectal contact situations. 'Shifting' or 'style-shifting' (Rickford and Eckert 2001) tend to be used more than CS to refer to the alternation between closely related varieties.

3. Sociolinguistic perspectives on code-switching

3.1. Macro-social perspectives

The first comprehensive sociolinguistic account of CS is attributed to Gumperz (Blom and Gumperz 1972; Gumperz 1982). His early accounts reflected Fishman's view that language use is stratified into (H)igh and (L)ow prestige social functions, which each call for specific varieties with a corresponding prestige index (Fishman 1972). Such an assumption of 'sociolinguistic indexicalities' (see Silverstein 2003) is particularly visible in the way Gumperz (1982) originally described 'we-codes' and 'they-codes'. The former, which index solidarity, typically coincide with L-varieties, whereas 'they-codes', which index distance, typically coincide with H-varieties. As a result, the alternation between we-codes and they-codes in conversation may convey an alternation between the two attitudes. The assumption of indexicality featured at the core of several sociolinguistic models of CS, before its systematic relevance to interpreting CS was called into question (Section 3.2).

The Markedness Model (MM; Myers-Scotton 1993b, 1998) posits sociolinguistic indexicality, distinguishing between 'unmarked codes' and 'marked codes' (varieties whose situational use reflects or breaks with macro-social conventions, respectively). The alternation between unmarked and marked codes in exchanges obeys Gricean maxims (Myers-Scotton's 'Rights and Obligations sets', or 'RO sets') which speakers, presented as rational actors, choose to implement (or not). Solidarity-marking codes (typically L-varieties) may be unmarked in specific social situations, and distance-marking ones (typically H-varieties) in others. The choice of a marked code over an unmarked one manifests the speaker's intention to renegotiate the RO set for the ongoing exchange into one compatible with the values of solidarity (or distance).

Communication Accommodation Theory (CAT; Gallois, Ogay, and Giles 2005) is another model that operates on the premise of sociolinguistic indexicality. It describes variation in terms of convergence and divergence, which can involve CS, and maintenance, which does not. These accommodation behaviors can be read as responses to, or attempts to negotiate, sociolinguistic differentials: 'upward' convergence of A with B, for example, is predictable when A has a lower social status than B, and will be effected through B's higher-status variety. Although it sees sociolinguistic indexicality in accommodation when sociolinguistic differentials are salient to the interactants, CAT also acknowledges that these behaviors can be psychologically motivated (when the interactants have affective bonds), or cognitively motivated (when one of the interactants' default variety is not necessarily intelligible to the other interactant).

Some macro-sociolinguistic perspectives on CS have focused on the grammatical forms of CS and what they see as their macrosocial correlates. Muysken (2000) proposed the most comprehensive macro-sociolinguistic typology of CS (see Section 2.1 for more details on types of CS), where insertional and alternational CS tend to occur in societal contexts marked by equal prestige across languages. By contrast, congruent lexicalization is characteristic of societal contexts in which one language holds more prestige than the other, such as in particular postcolonial settings where only the ex-colonizer's language is official, and indigenous languages are not (and are typically used in L-functions). Muysken leaves open whether sociolinguistic factors should be given more importance than linguistic ones in predicting what grammatical form CS is

likely to assume. Macrosocial explanations of grammatical diversity in CS are also found in Poplack (1987), who focuses on Canadian French-English and Spanish-English CS in the US, and in Treffers-Daller (1999), who compares Romance-Germanic CS in Strasbourg and Brussels.

Fewer sociolinguistic CS studies focus on social perceptions of CS. Those using matched-guise techniques overwhelmingly point towards a negative attitude towards CS, as exemplified by Chana and Romaine (1984) with Punjabi-English CS in the UK, Gibbons (1987) with Chinese-English CS in Hong Kong, and Lawson-Sako and Sachdev (1996) with French-Arabic CS in Tunisia. Other attitudinal studies within the ethnographic paradigm, such as Rampton (2014) among immigrant youths in London, or McCormick (2002) among Cape Town 'Coloureds', have identified covertly positive attitudes rooted in oppositional identities and resistance. Some sociolinguistic CS studies also deal with social variables impacting CS, including gender (which Poplack [1980] correlated with different frequencies and forms of CS in bilingual communities), as well as ethnicity (to some extent considered jointly with CS in studies of emergent urban ethnolects; see Rampton 2014).

3.2. Views of CS as a strategy

Sociolinguistics has increasingly focused on CS as an individual strategy deployed independently of macrosocial conventions. Much of that microsocial focus builds on Gumperz' ethnographic approach to CS, epitomized in his functional subdivision of CS (Blom and Gumperz 1972) into 'situational' and 'metaphorical' CS. In the former, the act of switching is triggered by external factors, such as changes in settings or in participant constellation. In the latter, by contrast, the speaker is the main actor; this type further involves the linguistic invocation of different personae, whose deployment need not be occasioned by situational changes. A function of metaphorical CS that sets it apart from sociolinguistically indexical types of CS is that of symbolic 'contextualization': individual code-switches may be used as conversational devices to cue new conversational footings, or frames, concepts originally elaborated by Goffman (1974).

Gumperz contributed the ethnographic notion that both analyst-oriented (or 'etic') and member-based (or 'emic') perspectives should direct interpretations of language variation. That led researchers to emphasize the potential disjunction between we-/they-codes and specific varieties. Swigart (1992) and Meeuwis and Blommaert (1998) found that languages were too mutually imbricated in their African CS data to signal an alternation between solidarity and distance. The observation that individual switches display varying degrees of pragmatic salience (and thus carry varying amounts of social meaning) led Auer (1999) to distinguish between 'language alternation', where individual switches possess salience, and 'language mixing', where they do not. 'Language mixing' may carry specific social meanings as a whole, and its recurrent observation has justified a radical rethinking of 'code' and CS.

Auer (1984) questioned whether CS really is an appropriate generic term to accommodate language mixing. The acknowledgement of language mixing as a system of its own has motivated a search for a more inclusive concept. One candidate has been the concept of 'repertoire', used in sociolinguistics to refer to the range of 'styles' (or 'registers' or 'genres') at the disposal of individual speakers (Rickford and Eckert 2001). Style

has been understood as a system of 'consistent' linguistic features, largely treated as synonymous with 'variety' in Labovian sociolinguistics, but more broadly defined by Irvine (2001: 23) as a "social semiosis of distinctiveness". Although Auer (2007) introduced the concept of 'style' under the purview of CS studies, the concept of 'style-shifting' seems to have been used mostly to describe the modes of alternation between related varieties of one language (see Section 2.2).

Blommaert (2010) introduced the 'super-diversity discourse' into sociolinguistics in combination with a distinctive sociolinguistic terminological framework centered around the concept of repertoire, set out mainly in Duarte and Gogolin (2013) and Arnaut et al. (2016). Originally coined by Vertovec (2007), super-diversity refers to the 'diversification of diversity' brought about by the accelerating pace of globalization. One linguistic consequence of super-diversity, in Blommaert's view, is the diversification of linguistic resources available in individual repertoires where the notion of speech community is becoming more elastic. This discourse uses specific terms to capture the linguistic 'mixedness' of contemporary linguistic phenomena, among them 'translanguaging' and 'polylanguaging'. Anchored in Swain's (2006) theory of language learning, these terms describe multilingual practices from the perspective of single system repertoires, within which there cannot be a question of discrete languages, as the term code-switching implies, following a general view that languages are essentially ideological constructs (Creese and Blackledge 2015; Jørgensen et al. 2011; see MacSwan [2017] for a critique).

Distinctive among CS studies, Auer (1984) argued for interpretations of CS within a theoretical framework based on ethnomethodology and CA, which focuses on conversational mechanisms. Leaning against this theoretical background, Auer proposes member-based, sequential interpretations of CS which reveal that some code-switches are endowed with conversation-structuring functions. Subsequent studies that refer to CA tend to weigh CA interpretations of CS against macrosocial ones (Auer 1998).

4. Acquisition and competence perspectives on code-switching

The study of how bilingual competence develops has mostly been associated with second language acquisition (SLA) research and bilingual first language acquisition (BFLA) (see Chapter 16). SLA originally bore the theoretical imprint of behaviorism by according much importance to the L1 in the L2 learning process. In the behaviorist view (Lado 1957), the L1 is present in L2 production in the form of 'transfers', which may be 'positive' (facilitating the replication of an L2 pattern), or 'negative' (leading to production errors, commonly referred to as 'interference'). The next influential theoretical framework in SLA was articulated around the Chomskyan assumption of an innate language building faculty, programmatically formulated by Selinker (1972) for his 'interlanguage hypothesis'. Adherents have characteristically focused on bilingual child language acquisition, and dwelt on errors not attributable to interferences, a term which Slobin (1985) has since replaced with the more neutral 'cross-linguistic influence'.

Under names such as 'code-mixing', 'mixing', or 'borrowing', CS has been discussed from a language acquisition perspective largely to determine how languages are simultaneously acquired in childhood (De Houwer 2009). Volterra and Taeschner (1978) contended that children exposed to two languages initially do not distinguish between them,

and accordingly engage in unconstrained CS. Despite occasional supporting evidence (Bader and Minnis 2000; Tracy 2000), the notion of a single system has been contested, based on indications that language separation and CS are subject to the parents' bilingual practices (Goodz 1994; Lanza 1997) and to the availability to children of translation equivalents (Deuchar and Quay 2000). Besides, Comeau, Genesee, and Lapaquette (2003) found that children interactionally adjust their CS patterns to those of their interlocutors, suggesting that bilingual children distinguish between their languages from early on.

Within the neuro-cognitive framework of language acquisition, Grosjean's (2001) model of bilingual language production had perhaps the strongest impact on CS typologies. His central notion is that of 'mode', itself rooted in the imagery of language activation: bilinguals can linguistically behave in 'monolingual mode', in which case they keep their languages separate in speech production, or in 'bilingual mode', in which case their languages alternate in production. Grosjean's notion of 'modes' underlies Myers-Scotton's MLF, in which the matrix language is presented as the more activated language (see Section 2.1). Myers-Scotton and Jake's (2016) subsequent 4-M model further elaborates the psycholinguistic underpinnings of the MLF by defining constraints to CS in relation to the differential levels of activation of the four categories of morphemes between which it distinguishes. Another attempt at incorporating psycholinguistic notions into the grammatical modelling of CS is subsumed in Clyne's (2003) 'triggering hypothesis', which posits that two, related languages can activate one another, producing congruent lexicalization (Section 2.1.).

Although psycholinguistic CS research traditionally rests on tasks administered within controlled laboratory settings (Gullberg, Indefrey, and Muysken 2009), some studies in the paradigm pay attention to the role played by interactional and social factors in the occurrence of CS. Along these lines, Kootstra, van Hell, and Dijkstra (2010) account for CS as an outcome of language co-activation facilitated by not only triggering effects produced by lexical and syntactic similarities across languages, but also by discourse-situational and socio-interactive factors (i.e. what they call 'interactive alignment'; see Branigan et al. 2007). Rather than looking at CS for the sake of determining the degrees of separation between bilingual children's languages, some BFLA studies have looked at it in conjunction with its developing socio-pragmatic functions, such as addressee-specification (Ervin-Tripp and Reyes 2005; Lanvers 2001), or turn-allocation (Cromdal 2001). BFLA research also has developed an ethnographic strand relying on long-term observation of classroom interactions, exemplified by Weber (2009) and Cheng (2003).

5. Conclusion

Much theoretical discussion on CS so far has centered on how much one should posit a surface or deep-level distinction between languages occurring in bilingual speech. Tentative answers have typically been forthcoming from discrete theoretical perspectives that rely on seemingly incompatible methodologies. Researchers in the field of CS studies should consider more attempts at cross-validation between its constitutive paradigms via multidisciplinary approaches to data collection and analysis.

6. References

Arnaut, Karel, Martha Sif Karrebæk, Massimiliano Spotti & Jan Blommaert (eds.)
2016 *Engaging superdiversity: Recombining spaces, times and language practices.* Bristol: Multilingual Matters.

Auer, Peter
1984 *Bilingual conversation*. Amsterdam: John Benjamins.

Auer, Peter
1998 Introduction: *Bilingual Conversation* revisited. In Peter Auer (ed.), *Code-switching in conversation: Language, interaction and identity*, 1–24. London: Routledge.

Auer, Peter
1999 From codeswitching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* 3(4). 309–332.

Auer, Peter (ed.)
2007 *Style and social identities: Alternative approaches to linguistic heterogeneity.* Berlin: Mouton de Gruyter.

Bader, Yousef & Denise D. Minnis
2000 Morphological and syntactic code-switching in the speech of an Arabic-English bilingual child. *Multilingua* 19(4). 383–403.

Bakker, Peter
1997 *A language of our own: The genesis of Michif, the mixed Cree-French language of the Canadian Métis*. New York: Oxford University Press.

Belazi, Hedi M., Edward J. Rubin & Almeida Jacqueline Toribio
1994 Code-switching and X-bar theory: The functional head constraint. *Linguistic Inquiry* 25(2). 221–237.

Blom, Jan-Petter & John J. Gumperz
1972 Social meaning in linguistic structures: Code-switching in Norway. In John J. Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics*, 407–434. New York: Holt.

Blommaert, Jan
2010 *The sociolinguistics of globalization.* Cambridge: Cambridge University Press.

Branigan, Holly P., Martin J. Pickering, Janet F. McLean & Alexandra A. Cleland
2007 Syntactic alignment and participant role in dialogue. *Cognition* 104(2). 163–197.

Chana, Urmi & Suzanne Romaine
1984 Evaluative reactions to Panjabi/English code-switching. *Journal of Multilingual and Multicultural Development* 5(6). 447–453.

Cheng, Karen Kow Yip
2003 Code-switching for a purpose: Focus on pre-school Malaysian children. *Multilingua* 22(1). 59–77.

Chomsky, Noam
1995 *The minimalist program*. Cambridge, MA: MIT Press.

Clyne, Michael
1987 Constraints on code switching: How universal are they? *Linguistics* 25(4). 739–764.

Clyne, Michael
2003 *Dynamics of language contact: English and immigrant languages*. Cambridge: Cambridge University Press.

Comeau, Liane, Fred Genesee & Lindsay Lapaquette
2003 The modelling hypothesis and child bilingual codemixing. *International Journal of Bilingualism* 7(2). 113–126.

Creese, Angela & Adrian Blackledge
2015 Translanguaging and identity in educational settings. *Annual Review of Applied Linguistics* 35. 20–35.

Cromdal, Jakob
2001 Overlap in bilingual play: Some implications of code-switching for overlap resolution. *Research on Language and Social Interaction* 34(4). 421–451.
de Bot, Kees
1992 A bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics* 13(1). 1–24.
De Houwer, Annick
2009 *Bilingual first language acquisition*. Clevedon: Multilingual Matters.
Deuchar, Margaret & Suzanne Quay
2000 *Bilingual acquisition: Theoretical implications of a case study*. Oxford: Oxford University Press.
Di Sciullo, Anne-Maria, Pieter Muysken & Rajendra Singh
1986 Government and code-mixing. *Journal of Linguistics* 22(1). 1–24.
Duarte, Joana & Ingrid Gogolin
2013 *Linguistic superdiversity in urban areas: Research approaches*. Amsterdam: Benjamins.
Ervin-Tripp, Susan & Iliana Reyes
2005 Child codeswitching and adult content contrasts. *International Journal of Bilingualism* 9(1). 85–102.
Fishman, Joshua A.
1972 *The sociology of language: An interdisciplinary social science approach to language in society*. Rowley: Newbury House.
Gallois, Cindy, Tania Ogay & Howard Giles
2005 Communication accommodation theory: A look back and a look ahead. In William Gudykunst (ed.), *Theorizing about intercultural communication*, 121–148. Thousand Oaks: SAGE.
Gibbons, John
1987 *Code-mixing and code-choice: A Hong Kong case study*. Clevedon: Multilingual Matters.
Goffman, Erving
1974 *Frame analysis*. New York: Harper & Row.
Goodz, Naomi S.
1994 Interaction between parents and children in bilingual families. In Fred Genesee (ed.), *Educating second language children*, 61–81. New York: Cambridge University Press.
Grosjean, François
2001 The bilingual's language modes. In Janet L. Nicol (ed.), *One mind, two languages: Bilingual language processing*, 1–23. Oxford: Blackwell.
Gullberg, Marianne, Peter Indefrey & Pieter Muysken
2009 Research techniques for the study of code-switching. In Barbara E. Bullock & Almeida J. Toribio (eds.), *The Cambridge handbook of code-switching*, 21–39. Cambridge: Cambridge University Press.
Gumperz, John. J.
1982 *Discourse strategies*. Cambridge: Cambridge University Press.
Hinskens, Frans, Peter Auer & Paul Kerswill
2005 The study of dialect convergence and divergence: Conceptual and methodological considerations. In Peter Auer, Frans Hinskens & Paul Kerswill (eds.), *Dialect change: Convergence and divergence in European languages*, 1–48. Cambridge: Cambridge University Press.
Irvine, Judith T.
2001 "Style" as distinctiveness: The culture and ideology of linguistic differentiation. In Penelope Eckert & John R. Rickford (eds.), *Style and sociolinguistic variation*, 21–43. Cambridge: Cambridge University Press.
Jørgensen, J. Normann, Martha S. Karrebæk, Lian M. Madsen & Janus S. Møller
2011 Polylanguaging in superdiversity. *Diversities* 14(2). 23–38.

Kootstra, Gerrit Jan, Janet G. van Hell & Ton Dijkstra
2010 Syntactic alignment and shared word order in code-switched sentence production: Evidence from bilingual monologue and dialogue. *Journal of Memory and Language* 63(2). 210–231.
Lado, Robert
1957 *Linguistics across cultures: Applied linguistics for language teachers.* Ann Arbor: University of Michigan Press.
Lanvers, Ursula
2001 Language alternation in infant bilinguals: A developmental approach to codeswitching. *International Journal of Bilingualism* 5(4). 437–464.
Lanza, Elizabeth
1997 *Language mixing in infant bilingualism: A sociolinguistic perspective*. Oxford: Clarendon Press.
Lawson-Sako, Sarah & Itesh Sachdev
1996 Ethnolinguistic communication in Tunisian streets. In Yasir Suleiman (ed.), *Language and ethnic identity in the Middle East and North Africa*, 61–79. Richmond: Curzon Press.
Le Page Robert B. & Andrée Tabouret-Keller
1985 *Acts of identity: Creole-based approaches to language and ethnicity.* Cambridge: Cambridge University Press.
MacSwan, Jeff
1999 *A minimalist approach to intrasentential code-switching*. New York: Garland.
MacSwan, Jeff
2005 Codeswitching and generative grammar: A critique of the MLF model and some remarks on 'modified minimalism'. *Bilingualism: Language and Cognition* 8(1). 1–22.
MacSwan, Jeff
2014 Programs and proposals in codeswitching research: Unconstraining theories of bilingual language mixing. In Jeff MacSwan (ed.), *Grammatical theory and bilingual codeswitching*, 1–34. Cambridge, MA: MIT Press.
MacSwan, Jeff
2017 A multilingual perspective on translanguaging. *American Educational Research Journal* 54(1). 167–201.
Mahootian, Shahrzad
1993 *A null theory of code switching*. Evanston, IL: Northwestern University dissertation.
McCormick, Kay
2002 *Language in Cape Town's District Six.* Oxford: Oxford University Press.
Meeuwis, Michael & Jan Blommaert
1998 A monolectal view of code-switching: Layered code-switching among Zairians. In Peter Auer (ed.), *Code-switching in conversation: Language, interaction and identity*, 76–98. London: Routledge.
Muysken, Pieter
1981 Halfway between Quechua and Spanish: The case for relexification. In Arnold Highfield & Albert Valdman (eds.), *Historicity and variation in creole studies*, 52–78. Ann Arbor: Karoma.
Muysken, Pieter
2000 *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.
Muysken, Pieter
2013 Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16(4). 1–22.
Myers-Scotton, Carol
1993a *Duelling languages: Grammatical structures in code-switching*. Oxford: Oxford University Press.

Myers-Scotton, Carol
1993b *Social motivations in code-switching: Evidence from Africa*. Oxford: Clarendon.
Myers-Scotton, Carol (ed.)
1998 *Codes and consequences: Choosing linguistic varieties*. Oxford: Oxford University Press.
Myers-Scotton, Carol
2002 *Contact Linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
Myers-Scotton, Carol & Janice L. Jake
2016 Revisiting the 4-M model: Codeswitching and morpheme election at the abstract level. *International Journal of Bilingualism* 21(3). 340–366.
Nortier, Jacomine
1990 *Dutch-Moroccan Arabic code switching*. Dordrecht: Foris.
Pfaff, Carol W.
1979 Constraints on language-mixing: Intrasentential code-switching and borrowing in Spanish-English. *Language* 55(2). 291–318.
Poplack, Shana
1980 Sometimes I'll start a sentence in English y termino en español: Toward a typology of code-switching. *Linguistics* 18(7–8). 581–618.
Poplack, Shana
1987 Contrasting patterns of code-switching in two communities. In Erling Wande, Jan Anward, Bengt Nordberg, Lars Steensland & Mats Thelander (eds.), *Aspects of multilingualism: proceedings from the fourth Nordic symposium on bilingualism, 1984*, 51–77. Uppsala: University of Uppsala.
Poplack, Shana, David Sankoff & Christopher Miller
1988 The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26(1). 47–104.
Poplack, Shana & Marjory Meechan
1995 Patterns of language mixture: nominal structure in Wolof-French and Fongbe-French bilingual discourse. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, 199–232. Cambridge: Cambridge University Press.
Rampton, Ben
2014 *Crossing: Language and ethnicity among adolescents*. London: Routledge.
Rickford, John R. & Penelope Eckert
2001 Introduction. In Penelope Eckert & John R. Rickford (eds.), *Style and sociolinguistic variation*, 1–18. Cambridge: Cambridge University Press.
Sankoff, David & Shana Poplack
1981 A formal grammar for code-switching. *Papers in Linguistics* 14(1). 3–45.
Selinker, Larry
1972 Interlanguage. *IRAL* 10. 209–231.
Silverstein, Michael
2003 Indexical order and the dialectics of sociolinguistic life. *Language and Communication* 23(3–4). 193–229.
Slobin, Dan I.
1985 Crosslinguistic evidence for the Language-Making Capacity. In Dan I. Slobin (ed.), *The crosslinguistic study of language acquisition: Vol. 2. Theoretical issues*, 1157–1256. Hillsdale: Erlbaum.
Swain, Merrill
2006 Languaging, agency and collaboration in advanced second language learning. In Heidi Byrnes (ed.), *Advanced language learning: The contributions of Halliday and Vigotsky*, 95–108. London: Continuum.

Swigart, Leigh
1992 Two codes or one? The insiders' view and the description of codeswitching in Dakar. *Journal of Multilingual and Multicultural Development* 13(1–2). 83–102.
Thomason, Sarah G.
2001 *Language contact: An introduction*. Washington D.C.: Georgetown University Press.
Thomason, Sarah G. & Terrence Kaufman
1988 *Language contact, creolization and genetic linguistics.* Berkeley, CA: University of California Press.
Tracy, Rosemarie
2000 Language mixing as a challenge for linguistics. In Susanne Döpke (ed.), *Cross-linguistic structures in simultaneous bilingualism*, 11–36. Amsterdam: Benjamins.
Treffers-Daller, Jeanine
1999 Borrowing and shift-induced interference: Contrasting patterns in French-Germanic contact in Brussels and Strasbourg. *Bilingualism: Language and Cognition* 2(1). 1–22.
Treffers-Daller, Jeanine
2009 Code-switching and transfer: An exploration of similarities and differences. In Barbara E. Bullock & Almeida Jacqueline Torribio (eds.), *The Cambridge handbook of linguistic code-switching*, 58–74. Cambridge: Cambridge University Press.
Trudgill, Peter
2006 *New-dialect formation: The inevitability of colonial Englishes*. Edinburgh: Edinburgh University Press.
van Coetsem, Frans
1988 *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
Vertovec, Steven
2007 Super-diversity and its implications. *Journal of Ethnic and Racial Studies* 30(6). 1024–1054.
Vogt, Hans
1954 Language contacts. *Word* 10(2–3). 365–374.
Volterra, Virginia & Troute Taeschner
1978 The acquisition and development of language by bilingual children. *Journal of Child Language* 5(2). 311–326.
Weber, Jean-Jacques
2009 *Multilingualism, education and change*. Frankfurt am Main: Peter Lang.
Weinreich, Uriel
1953 *Languages in contact: Findings and problems* (= Publications of the Linguistic Circle of New York 1). New York: Linguistic Circle of New York.
Woolford, Ellen
1983 Bilingual code-switching and syntactic theory. *Linguistic Inquiry* 14(3). 520–536.

Gerald Stell, Hong Kong

15. Language contact across the lifespan

1. Introduction
2. Children
3. Adolescents
4. Adults
5. Older speakers
6. Future directions
7. References

1. Introduction

The outcome of a multilingual encounter depends on a number of factors. Contact may be of various kinds and may differ in duration, and speakers may belong to different sociocultural groups depending on such variables as socioeconomic status, gender, and, crucially, *age*. Furthermore, the outcomes may consequently differ in nature − from lexical loan words, to prosodic transfer and morphosyntactic simplification, to the formation of new contact varieties. This chapter provides a state of the art overview of the ways in which age affects the outcome of a linguistic contact situation, and the way in which age is relevant to the formation of that outcome.

Chronological age in industrialized society serves as the indication of a person's place in society and in the life cycle, measured in years passed since birth. Such a biological or chronological approach is not, however, typical of all societies. The literature on the topic of age has acknowledged this by shifting from viewing age as solely chronological and biological to viewing it as a social construct, situated within frameworks that view language as critical to shaping and creating our social worlds (Andrew 2012). Framing age as social renders it compatible with recent dynamic approaches to identity: age, as an act of identity, is performed, rather than a static state. Age, then, can be approached from different perspectives, including four possible vantage points: biological age, psychological age, functional age, and ultimately, social age. It can also be viewed as all of the above (Cameron 2011: 208).

Approaching language contact studies from the perspective of the individual's age is challenging because most approaches differ as to whether they focus on the process, or practice, of multilingual encounters, or on the outcome, or product, of such encounters. It is rare to find studies that explicitly consider both − and even rarer to find such studies that consider age as a variable. As we return to below, notable exceptions include Kerswill and Williams (2000), Trudgill (2011), and Matras (2009). Auer (2014: 295) also points to the necessity of shifting focus from outcome (e.g. a mixed language) to process (e.g. the practice of extensive borrowing).

Defining contact-induced change as "any linguistic change that would have been less likely to occur outside a particular contact situation", Thomason (2001: 61−63) illustrates this focus on the diachronic and structural dimensions of language contact, and less on the individuals involved in multilingual practices that actually enable these changes. A case in point is Heine and Kuteva's (2005: 116) discussion of contact-induced grammaticalization, in which they state that they will have "little to say about the process leading to the product since it is still in the main poorly understood". As linguistic change is often observed post hoc (Wagner 2012: 371−372), it is particularly difficult to assume anything about the nature of the interaction of the speakers that created this change.

https://doi.org/10.1515/9783110435351-015

Many historical linguists avoid it altogether, creating a lack of records of historical change that take context and/or speaker type into consideration. A different approach involves a focus on the study of linguistic interactions that may create contact-induced change (e.g. code-switching), and hence a focus on the process of multilingual practices (Treffers-Daller and Mougeon 2005).

Despite generating a "mild interest", as noted by Eckert (1996: 167), and not having been explicitly studied as a sociolinguistic variable, age as a sociolinguistic dimension is routinely included in, for example, variationist studies, alongside variables such as gender and ethnicity (Andrew 2012: 39). The inclusion of age in such studies is typically to uncover or illustrate linguistic change, with age serving as a reflection of historical change. Age can reflect a shift in language across a speech community (reflecting historical change), or a shift in language in the individual (*age grading*). From the former perspective, older speakers provide a mirror into the language of older times, through a methodology referred to as the *apparent time construct*, i.e. using the present to explain the past (Labov 2006; see also Bailey 2008). Despite being widely applied, this approach is also recognized for having shortcomings, and it has been noted that this approach actually underestimates change (Sankoff and Blondeau 2007).

With an increased focus on the multilingual individual, there is a growing awareness that contact is indeed a metaphor. What is in contact are people speaking different languages. Although this awareness began with Weinreich's (1953) iconic study, the actor − or speaker − is in fact absent from a majority of studies on the outcomes of language change, and language contact in general. Various waves of linguistic research have contributed to an increased focus on the individual in contact linguistics, but this increased awareness is probably symptomatic of a general epistemological change that involves an increased interest in the multilingual individual. This is a change that involves a shift from more structure-based to more agency-based theories of language, or from macro-oriented to micro-oriented approaches, as well as an alignment between sociolinguistics and linguistic anthropology (Kramsch 2009). The growing understanding of multilingual practice has also enabled the view that the languages − or repertoires − of multilingual speakers interact in harmony (Busch 2012). In other words, in sociolinguistics there is currently a move away from considering 'named languages' being in contact, toward a view that takes as the point of departure the multilingual individual's linguistic repertoire, which may be composed of elements from various 'named languages' (Canagarajah 2013; Makoni and Pennycook 2007).

In this chapter, we focus on types of interactions between speakers that belong to different age groups − interactions that may (or may not) in turn lead to the above-mentioned structural changes. We focus on how multilingual speakers dynamically switch between the languages in which they are competent (code-switching, language mixing), or utilize various elements in their linguistic repertoires (translanguaging). In the following sections, we address four different age groups (children, adolescents, adults, and older speakers) in light of language contact, with each section presenting the perspective of psycholinguistic versus sociolinguistic approaches.

2. Children

The role of children in studies of language contact has typically been examined from the perspective of acquisition when considering preschool children − whether the case

is one of bilingual first language acquisition or second language acquisition. Age has been a critical variable in studies of bilingual first language acquisition, with exposure to both languages already from birth being a stringent criterion for evaluating linguistic output from children, discussed in depth in Chapter 16 (De Houwer 2009). In more recent studies of young bilingual children, language contact per se has received less attention, while issues concerning input and exposure are highlighted (Carroll 2017; Grüter and Paradis 2014). Language contact in bilingual and multilingual children who have their first exposure to another language after their first chronological language(s) has been studied from both an acquisition point of view, and from a language use perspective, i.e. both psycholinguistically and sociolinguistically. How long childhood lasts is indeed to a large extent culturally determined. However, as noted above, in the extant studies of language contact and children, the age variable is critical for demarcating multiple first language acquisition from early second language acquisition. Furthermore, the language of preschool children, usually pre-literate, is distinguished from that of school-age children before puberty.

2.1. Psycholinguistic aspects

Earlier studies of bilingual first language acquisition promoted a one-system hypothesis, and posited that language contact was a sign of a lack of language differentiation by the young child acquiring more than one language simultaneously (Vihman 1985; Volterra and Taeschner 1978; see Lanza [2004] for a review). However, newer research offers robust evidence of the child's capacity to separate her/his languages, indeed from a very early age. In fact, infant perception studies indicate early perception of language-specific features (Fennell, Tsui, and Hudon 2016). Nonetheless, issues of language dominance are also relevant to assessing language contact in the speech of young language acquirers (Yip and Matthews 2007). Interactional analyses, furthermore, illustrate how the young bilingual child differentiates between her/his languages and shows sensitivity towards social parameters of language use, and hence that s/he can indeed code-switch and use basic structural language contact patterns attested in the speech of older speakers in the code-switching literature (Lanza 1997, 2004; Lanza and Li Wei forthcoming; Cantone 2007).

In studies of second language acquisition, age has been an important construct, particularly due to the critical period hypothesis, which states that we are biologically predestined to learn language with ease only before a certain age (around puberty), rendering age a determining factor in second language learning (Long 1990). Descriptions of multilingual practices are abundant, and de Bot and Makoni (2005: 10) note that studies of multilingual practices within the field of applied linguistics have given extensive attention to the age factor, mainly from the perspective of early L1/L2 acquisition and whether native-like proficiency is possible to achieve after puberty.

The degree of language *attrition,* or individual-level language loss in healthy speakers, varies dramatically, depending on the age at which the speakers became exposed to their new linguistic environments. The key moment here is puberty. If exposure to the new linguistic environment takes place before the onset of puberty, attrition is more severe. Schmitt (2006, 2010), for example, finds that Russian emigrants to the US were unable to accurately produce oblique cases (although their proficiency in the nominative

persisted). Furthermore, Schmitt (2006, 2010) found that in children, there is more evidence for both retention and substitution of L1 case assignments than in adults − more evidence for retention and substitution than for loss. Note that child speakers who have attained school age, but not puberty, will generally be able to use most grammatical features of their L1 in a target-like way. Nonetheless their language remains vulnerable to attrition. This is the reason why age of migration is crucial to the extent to which they experience language attrition.

2.2. Sociolinguistic aspects

More recent attention to the young schoolchild's use of more than one language has focused on what is referred to as *translanguaging* (García and Li Wei 2014), a term first used in the context of the classroom (see also Chapters 14 and 34). Teachers allowing their students to use the full range of their linguistic repertoires, or translanguaging, contribute to the multilingual child's learning.

Traditionally understood as being a phenomenon only of young adults, *multiethnolects* (see Section 3.2), or non-standard 'mixed' varieties used by speakers with different linguistic backgrounds, are also used by children. In a study on language contact in London, Cheshire et al. (2011) find pronunciation patterns in children as young as five that are compatible with patterns found in surrounding multiethnolects. Interestingly, these patterns are not compatible with the parent's English pronunciation, and the authors conclude that this is likely so because the children do not speak English with their parents, and therefore are exposed to English primarily through interaction with other multilingual children in 'shift-induced interference' (Thomason and Kaufman 1988). The same was found for 8-year-olds in London, who mainly made use of the multicultural London vowel system. Together, these findings indicate that features from multiethnolects associated with youth language are not only used by young adults, but also other age groups.

Kerswill et al. (2013) provide a detailed account of the role of age groups in contact situations, and in the development of the English language specifically. The authors suggest a tripartite division to assess the speech of children and adolescents. Infants and children (0−6), they argue, are special, in the sense that they inherit a substantial sociolinguistic competence from their parents or caregivers. Preadolescents (6−12) find themselves in the specific situation in which other peers gradually become the primary influence on their dialect and contact feature acquisition, while adolescents (12−17), on the other hand, demonstrate a growing peer-group involvement, as we shall see in the next section.

3. Adolescents

Adolescents have received substantial attention in the literature on language contact, linguistic innovation, and change (Cutler and Røyneland 2018). They are seen as prime innovators and users of non-standard speech forms (Eckert 1988, 2003; Nortier and Svendsen 2015), until this period ends in the late teens. Adolescents incrementally in-

crease their use of new norms, but this increase levels off at about the age of 16/17, a point which is referred to as the *adolescent peak* (Kirkham and Moore 2013). Adolescence differs from the other age groups discussed here, in that as an age group, it is relatively new in the research literature. Adolescence being seen as a separate age group is largely a product of industrialized society, more specifically of its institutionalized secondary education, which has resulted in segregated, age-homogeneous groups involving adolescents only. In these networks, besides the construction of identities, linguistic innovation and creativity thrives (Eckert 2003), due to the flexibility of language norms and the high tolerance for linguistic variation, and the socio-cognitive period of identity searching that forms linguistic development (more than the stabilization of a grammar, as in children). Different approaches to age thus have sociolinguistic consequences. In societies in which adolescents are expected to work, for example, they do not form age-homogeneous groups similar to those found in current Western societies' systems of secondary education, where linguistic innovation flourishes.

3.1. Psycholinguistic aspects

Adolescence is crucial for understanding attrition (see also Chapter 17). Here, age of emigration is critical (Schmid 2011): if you emigrate after puberty, you are less likely to experience language attrition. In attrition, many elements come into play, such as age of emigration, age since emigration, and age at the time of the investigation. Studies of attrition occur most effectively with individuals after the onset of puberty, because attrition can be confused with incomplete acquisition in the case of individuals who emigrated during childhood. Once a speaker has acquired knowledge of a second language, there will necessarily be some interaction not only from the first to the second language, but also from the second language to the first (in cases involving two languages). This view presupposes the hypothesis of a critical period (which has been heavily criticized, see Birdsong [1999] for a much-cited overview; see also Hartshorne, Tenenbaum, and Pinker [2018] for a discussion on its age of offset), with some studies finding that the first language can deteriorate, and sometimes even completely disappear, if emigration takes place before the age of 12. It appears that even though second language learning is easier when engaged in before the critical period is over, it is also more likely that the first language will be subject to more radical attrition at this very same age.

Adolescent multilingual competence also plays a key role in studies of simplification due to post-threshold second language learning (McWhorther 2004; Trudgill 2011). This view entails that adolescents have passed the critical period of language learning, and therefore will fail to learn structures that are 'L2-difficult'. This inability might lead to simplification of a morphological, syntactic, and phonological nature (for a discussion of a possible contact-induced simplification in Spanish spoken in Buenos Aires, Argentina, see Fløgstad [2016: 179–185]). This assumption is far-reaching. Pavlenko (2014: 38–39), for example, argues that the assertion that languages are by definition equally complex is meaningless, because languages spoken in industrialized, contemporary societies necessarily undergo simplification because they are learned by a number of speakers who have acquired them as a second language. Since this is not the case for languages spoken in small, economically self-sufficient groups, the latter languages tend to display greater morphosyntactic complexity.

3.2. Sociolinguistic aspects

As we observed in 2.2, multiethnolects, or urban youth speech styles, are typically associated with youth and young adults. A multiethnolect is used to refer to different linguistic varieties: mixed speech of speakers with different heritage languages, with the major dominant language, or with an ethnically neutral variety, which nevertheless includes non-standard and innovative phonetic, grammatical, and discourse-pragmatic features (see Nortier and Svendsen [2015] on urban youth styles in Scandinavia and Rampton's [1995] classical work on crossing). Both urban youth styles and crossing are typically associated with adolescent speakers, and are often assumed not to continue beyond adolescence (see, however, Section 4). Cheshire et al. (2011), in a study of multiethnolects in London, treat the formation of multicultural urban vernaculars in detail. They argue that for the London case, the nature of the multiethnolect may have depended on its formation. In certain areas, due to the lack of native speaker models, large groups acquire the majority language from other second language speakers, a phenomenon sometimes referred to as 'group second language acquisition' and shift-induced interference. Given the lack of native input, the version of the majority language acquired by these speakers may show signs of, for example, simplification. If the minority group is well-integrated, these simplifications may spread to the speakers of the majority language. Note, however, that Cheshire et al. (2011) find no evidence of imperfect learning in their study.

4. Adults

If adolescents are the innovators, then adults are seen as conservative, possibly to comply with the pressure to use standard forms of language at the workplace that begins to occur after their late teens. Compared to children and adolescents, adults have relatively stabilized grammars, and bring these with them into new interactions, but see Grosjean and Py (1991), which shows that even adults' first language can change in a contact situation. Only adults are usually portrayed in the literature as the 'default', 'unmarked' age group in language use. Historically, adults have been the main subject of research for methodological reasons: age as a variable is readily applied to contemporary sources, but less so to historical sources, as studies that are based on written language necessarily reflect the language of those who knew how to write − often adults.

4.1. Psycholinguistic aspects

Adults are often thought of as having grammars that represent an 'endpoint', and their linguistic competence is treated without a developmental perspective (a perspective that involves children as acquirers, adolescents as innovators, and older speakers as experiencing loss in linguistic and cognitive abilities). As such, studies involving adults rarely involve fine-grained distinctions, and typically treat adults as a homogeneous group (de Bot and Makoni 2005). Adults are rarely distinguished in terms of being split into subgroups, and are rarely viewed as interesting per se, although age of second language acquisition can be an important variable in psycholinguistic testing of the effects of

bilingualism, that is, whether the adult learned the second language as a child or later in life (Ortega 2013). Typically, no distinctions are made between, for example, 30-, 40-, or 50-year old speakers in the second language acquisition literature (see Andrew [2012: 15–19] for an overview), but such nuances are included in studies of errors in elderly monolingual speakers; Burke and Shafto (2004), for example, find that old adults make more production errors than young adults.

4.2. Sociolinguistic aspects

When adults do contribute to the formation of a new dialect or variety, it is often assumed that their contribution will create simplifications. This is because their acquisition will involve problems with learning the difficult, or marked, structures of the new variety, as adults have passed the critical threshold for language acquisition. Still, this does not mean that adults do not change their languages at all – indeed, their ability to do so is emphasized within cognitive and usage-based approaches to language (Bybee 2010), as well as Dynamic Systems Theory (de Bot et al. 2013). Labov (2007) notes that adults participate in ongoing changes, albeit more sporadically, and at a slower rate, than children. He also finds that adult acquisition leads to the loss of the "fine structure of the linguistic system being transmitted" (Labov 2007: 380), compatible with findings emphasizing that post-threshold second language learning leads to simplification (Trudgill 2011). Note that this assertion has been challenged: Kerswill and Williams (2000) argue that by viewing scenarios that involve post-threshold second language learners as creators of simplification, we are ignoring important distinctions. For example, few migrations or settlements involve only adults; children from these groups will rapidly form a new speech community, and their exposure to their native languages will depend on a variety of linguistic and demographic factors.

Traditionally, multiethnolects were considered to be used by young adults only. However, we now know that such hybrid styles may persist. Sheng, for example, began as a typical urban youth variety in Nairobi, Kenya, but is now acquired as a first language (Dorleijn, Mous, and Nortier 2015). There is increasing evidence that multiethnolects may survive beyond adulthood. Rampton (2015) discusses the case of a British-born executive of Pakistani descent, with long-lasting relationships with speakers of many other languages, and finds a hybrid speech style that appears to be stable in his repertoire. He also proposes the term 'contemporary urban vernacular' to account for the fact that these varieties – sometimes referred to as 'urban youth speech styles' – are not restricted to young speakers only. Although there is evidence for the use of multiethnolects in adults, it continues to be a phenomenon primarily associated with young adulthood. It is also notably persistent: young people would use such styles 20 years ago, and similar structures are found in young adults today.

5. Older speakers

After decades of neglect (cf. the so-called 'gerontological lament' [Coupland, Coupland, and Giles 1991: 9]), there is growing research on multilingualism in the older population,

mainly as examples of an individual whose linguistic competence is in decline (Plejert, Lindholm, and Schrauf 2017). The mere definition of 'old' is not fixed to one chronological age, but is typically associated with loss of societal roles as well as cognitive decline. However, studies on the older generation's linguistic behavior have yet to be conducted within the field of 'language contact and its consequences'. Older speakers are typically included not due to an interest in their language per se, but either as representatives of older times (in apparent time studies), or as examples of post-threshold migrants (in attrition studies). The former appeals to historical causes to explain elders' speech, as opposed to alluding to inherent or environmental causes (Coupland, Coupland, and Giles 1991: 5–7). This group has traditionally had the position of being the stable, static representation of older times, and is rarely distinguished in terms of being split into subgroups. Exceptions include a division into young/old (64–76) and old/old (77 and up) (see Coupland, Coupland, and Giles 1991: 9; Gitterman, Goral, and Obler 2012; de Bot and Makoni 2005).

5.1. Psycholinguistic aspects

As for the psycholinguistic aspects concerning older speakers, it is true that most participants in studies of language attrition, for example, are relatively old – usually in their sixties or older. This, however, is usually not because the speaker's age is seen as interesting in itself. Investigators usually prefer participants that emigrated after puberty, but who have stayed in the country for a substantial amount of time. In addition, older speakers are often recruited for practical reasons: many older speakers are able to dedicate time to time-consuming projects, and also find themselves at a period in their lives in which they feel nostalgic about their country of origin (Schmid 2011). Older speakers are often also thought to experience language reversal, or decay in proficiency, in the second language. However, as Schmid (2011) notes, such claims are typically unsupported myths, and the few large scale studies that actually do address the issue find no evidence for such second language decay. There is also some evidence that older speakers become less conservative in their speech style with age. As Eckert (1996) notes, this is perhaps associated with the reduced burden associated with childcare and pressure to comply with workplace norms. With certain exceptions, it is a fact that studies of older speakers that do not have the perspective of decline are rare, but there is a growing literature in the study of multilingual dementia (e.g. Knoph, Simonsen, and Lind 2017).

5.2. Sociolinguistic aspects

As mentioned initially, older speakers are typically included in studies of language change not because of an interest in their talk per se, but as representative of older times in apparent time studies. Similarly, older speakers may represent a window into otherwise dying languages, and proficiency in a language by older members of a community is a sign that the language is threatened.

The concept of L1-regression – reverting to a first language in a second language setting – has been explained by environmental causes. A growing population of multi-

lingual older speakers entails a whole range of new challenges. Nursing homes, for example, provide an illustration of this, where older speakers with migrant backgrounds may have problems communicating with staff because they have reverted to their first language, and speakers of the majority language may have trouble communicating with staff who do not speak the local language (de Bot and Makoni 2005: 23–24). Both scenarios illustrate less-than-optimal types of interaction. There is less code-switching in older speakers, because they participate in less differentiated social networks that would require language mixing, and studies have shown that older speakers code-switch less than adolescents (David Maya et al. 2009).

6. Future directions

Although the actor-centered approach to language change and contact has allowed for substantial advances within the field of language contact, there is an obvious need for further implementation of the individual within this field, especially regarding the combination of process-centered approaches (such as code-switching) with product-centered approaches (those investigating the structural outcomes of the multilingual encounters) – studying contact-induced change *in medias res*. Sankoff (2006) notes that the greatest potential for developing this field is through the reintegration of the individual into the overall matrix of the speech community, and that this approach represents the greatest challenge to, and the greatest scope in, contact linguistics research. This assertion is still relevant, and only when such an integration takes place, and when focus is on the individual as the real locus of language change will age as a variable find its suitable place in studies of language contact. Large-scale studies that trace the individual's social networks (Petré 2017) are a step in that direction, and including the same methodology in multilingual settings would likely yield interesting results.

Furthermore, two age groups stand out as groups in particular need of more research. First, we see the language of adults as deserving to be studied in its own right, beyond being viewed as a static entity. How much innovation is possible in the adult linguistic system? Are monolingual adults capable of making substantial changes to their grammars and sound systems (as is argued by proponents of cognitive linguistics [Bybee 2010], but typically dismissed by proponents of generative approaches [Kroch 2008])? All in all, more fine-grained analyses of adults, possibly dividing adults into smaller age cohorts or by life event, not chronology, would be welcome. Second, an increased focus on the elderly is necessary in order to understand more about language contact, learning, and attrition in this group (Schmid 2011). Together, the need for more research on adults and the elderly confirms the need for an increased focus on the lifespan perspective.

7. References

Andrew, Patricia
2012 *The social construction of age: Adult foreign language learners. Second language acquisition*. Bristol: Multilingual Matters.

Auer, Peter
2014 Language mixing and language fusion: When bilingual talk becomes monolingual. In Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder & Achim Rabus (eds.), *Congruence in contact-induced language change: Language families, typological resemblance, and perceived similarity*, 294–334. Berlin/Boston: De Gruyter.
Bailey, Guy
2008 Real and apparent time. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 312–331. Oxford: Blackwell.
Birdsong, David
1999 *Second language acquisition and the critical period hypothesis: Second language acquisition research*. London: Erlbaum.
Burke, Deborah M. & Meredith A. Shafto
2004 Aging and language production. *Current Directions in Psychological Science* 13(1). 21–24.
Busch, Brigitta
2012 The linguistic repertoire revisited. *Applied Linguistics* 33(5). 503–523.
Bybee, Joan
2010 *Language, usage, and cognition*. Cambridge: Cambridge University Press.
Cameron, Richard
2011 Aging, age, and sociolinguistics. In Manuel Díaz-Campos (ed.), *The handbook of Hispanic sociolinguistics*, 207–229. Malden, MA: Wiley Blackwell.
Canagarajah, A. Suresh
2013 *Translingual practice: Global Englishes and cosmopolitan relations*. London: Routledge.
Cantone, Katja
2007 *Code-switching in bilingual children*. Dordrecht: Springer Netherlands, Dordrecht.
Carroll, Susanne E.
2017 Exposure and input in bilingual development. *Bilingualism: Language and Cognition* 20(1). 3–16.
Cheshire, Jenny, Paul Kerswill, Susan Fox & Eivind Torgersen
2011 Contact, the feature pool and the speech community. *Journal of Sociolinguistics* 15(2). 151–196.
Coupland, Nikolas, Justine Coupland & Howard Giles (eds.)
1991 *Language, society and the elderly: Discourse, identity and ageing*. Oxford: B. Blackwell.
Cutler, Cecelia & Unn Røyneland
2018 *Multilingual youth practices in computer mediated communication.* Cambridge: Cambridge University Press.
David Maya, Khemlani, Mei Tien Wendy Yee, Yeok Meng Ngeow & Kah Hui Gan
2009 Language choice and code switching of the elderly and the youth. *International Journal of the Sociology of Language* 200. 49–74.
de Bot, Kees & Sinfree Makoni
2005 *Language and aging in multilingual contexts*. Clevedon: Multilingual Matters.
de Bot, Kees, Wander Lowie, Steven L. Thorne & Marjolijn Verspoor
2013 Dynamic systems theory as a comprehensive theory of second language development. In María del Pilar García Mayo, María Juncal Gutiérrez Mangado & María Martínez-Adrián (eds.). *Contemporary perspectives on second language acquisition*. 167–189. Amsterdam: John Benjamins Publishers.
De Houwer, Annick
2009 *Bilingual first language acquisition*. Bristol: Multilingual Matters.

Dorleijn, Margreet, Maarten Mous & Jacomine Nortier
2015 Urban youth speech styles in Kenya and the Netherlands. In Jacomine Nortier & Bente A. Svendsen (eds.), *Language, youth, and identity in the 21st century: Linguistic practices across urban spaces*, 271–289. Oxford: Oxford University Press.
Eckert, Penelope
1988 Adolescent social structure and the spread of linguistic change. *Language in Society* 17(2). 183–207.
Eckert, Penelope
1996 Age as a sociolinguistic variable. In Florian Coulmas (ed.), *The handbook of sociolinguistics*, 151–167. Oxford: Blackwell.
Eckert, Penelope
2003 Language and adolescent peer groups. *Journal of Language and Social Psychology* 22(1). 112–118.
Fennell, Christopher T., Angeline Sin-Mei Tsui & Tamara M. Hudon
2016 Speech perception in simultaneously bilingual infants. In Elena Nicoladis & Simona Montanari (eds.), *Bilingualism across the lifespan: Factors moderating language proficiency*, 43–62. Washington, DC: American Psychological Association.
Fløgstad, Guro
2016 *Preterit expansion and perfect demise in Porteño Spanish and beyond: A critical perspective on cognitive grammaticalization theory.* Leiden/Boston: Brill.
García, Ofelia & Li Wei
2014 *Translanguaging: Language, bilingualism and education*. Basingstoke: Palgrave Macmillian.
Gitterman, Martin R., Mira Goral & Loraine K. Obler
2012 *Aspects of multilingual aphasia*. Bristol: Multilingual Matters.
Grosjean, François & Bernard Py
1991 La restructuration d'une première langue: L'intégration de variantes de contact dans la compétence de migrants bilingues. *La Linguistique* 27(2). 35–60.
Grüter, Theres & Johanne Paradis (eds.)
2014 *Input and experience in bilingual development*. Amsterdam/Philadelphia: John Benjamins.
Hartshorne, Joshua K., Joshua B. Tenenbaum & Steven Pinker
2018 A critical period for second language acquisition. Evidence from 2/3 million English speakers. *Cognition* 177. 263–277.
Heine, Bernd & Tania Kuteva
2005 *Language contact and grammatical change*. Cambridge: Cambridge University Press.
Kerswill, Paul & Ann Williams
2000 Creating a new town koine: Children and language change in Milton Keynes. *Language in Society* 29(1). 65–115.
Kerswill, Paul, Jenny Cheshire, Susan Fox, Eivind Torgersen, Daniel Schreier & Marianne Hundt
2013 *English as a contact language: The role of children and adolescents*. Cambridge: Cambridge University Press.
Kirkham, Sam & Emma Moore
2013 Adolescence. In Jack K. Chambers & Natalie Schilling (eds.), *The handbook of language variation and change*, 277–296. Oxford: Wiley-Blackwell.
Knoph, Monica I. Norvik, Hanne Gram Simonsen & Marianne Lind
2017 Cross-linguistic transfer effects of verb-production therapy in two cases of multilingual aphasia. *Aphasiology* 31(12). 1482–1509.
Kramsch, Claire
2009 *The multilingual subject: What foreign language learners say about their experience and why it matters*. Oxford: Oxford University Press.

Kroch, Anthony
2008 Syntactic change. In Mark Baltin & Chris Collins (eds.). *The handbook of contemporary syntactic theory*, 699–729. Malden/Oxford: Wiley Blackwell.
Labov, William
2006 *The social stratification of English in New York City*, 2nd edn. Cambridge: Cambridge University Press.
Labov, William
2007 Transmission and diffusion. *Language* 83(2). 344–387.
Lanza, Elizabeth
1997 Language contact in bilingual two-year-olds and code-switching: Language encounters of a different kind? *International Journal of Bilingualism* 1(2). 135–162.
Lanza, Elizabeth
2004 *Language mixing in infant bilingualism: A sociolinguistic perspective*. Oxford: Oxford University Press.
Lanza, Elizabeth & Li Wei
Forthcoming Code-switching and children. *Oxford bibliographies online*. Oxford: Oxford University Press.
Long, Michael H.
1990 The least a second language theory needs to explain. *Tesol Quarterly* 24(4). 649–666.
Makoni, Sinfree & Alastair Pennycook (eds.)
2007 *Disinventing and reconstituting languages*. Clevedon: Multilingual Matters.
Matras, Yaron
2009 *Language contact*. Cambridge/New York: Cambridge University Press.
McWhorter, John
2004 *Defining creole*. New York: Oxford University Press.
Nortier, Jacomine & Bente Ailin Svendsen (eds.)
2015 *Language, youth and identity in the 21st century: Linguistic practices across urban spaces*. Cambridge: Cambridge University Press.
Ortega, Lourdes
2013 *Understanding second language acquisition*. London: Routledge.
Pavlenko, Aneta
2014 *The bilingual mind: And what it tells us about language and thought*. Cambridge: Cambridge University Press.
Petré, Peter
2017 The extravagant progressive: an experimental corpus study on the history of emphatic [BE Ving]. *English Language & Linguistics* 21(2). 227–250.
Plejert, Charlotta, Camilla Lindholm & Robert W. Schrauf
2017 *Multilingual interaction and dementia*. Bristol: Multilingual Matters.
Rampton, Ben
1995 *Crossing: Language and ethnicity among adolescents*. London: Longman.
Rampton, Ben
2015 Contemporary urban vernaculars. In Jacomine Nortier & Bente A. Svendsen (eds.), Language, youth and identity in the 21st century: Linguistic practices across urban spaces, 24–44. Cambridge: Cambridge University Press.
Sankoff, Gillian
2006 Age: Apparent time and real time. In Keith Brown (ed.), *Elsevier encyclopedia of language and linguistics*, 110–116. Amsterdam: Elsevier.
Sankoff, Gillian & Hélène Blondeau
2007 Language change across the lifespan: /r/ in Montreal French. *Language* 83(3). 560–588.
Schmid, Monika S.
2011 *Language attrition*. Cambridge: Cambridge University Press.

Schmitt, Elena
2006 The 'bare bones' of language attrition. *Journal of Slavic Linguistics* 14(2). 263–288.
Schmitt, Elena
2010 When boundaries are crossed: Evaluating language attrition data from two perspectives. *Bilingualism: Language and Cognition* 13(1). 63–72.
Thomason, Sarah G.
2001 *Language contact*. Edinburgh: Edinburgh University Press.
Thomason, Sarah G. & Terrence Kaufman
1988 *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
Treffers-Daller, Jeanine & Raymond Mougeon
2005 The role of transfer in language variation and change: Evidence from contact varieties of French. *Bilingualism: Language and Cognition* 8(2). 93–98.
Trudgill, Peter
2011 *Sociolinguistic typology: Social determinants of linguistic complexity.* Oxford: Oxford University Press.
Vihman, Marilyn
1985 Language differentation by the bilingual infant. *Journal of Child Language* 12(2). 297–324.
Volterra, Virginia & Traute Taeschner
1978 The acquisition and development of language by bilingual children. *Journal of Child Language* 5(2). 311–326.
Wagner, Suzanne Evans
2012 Age grading in sociolinguistic theory. *Language and Linguistics Compass* 6(6). 371–382.
Weinreich, Uriel
1953 *Languages in contact.* Mouton: The Hague.
Yip, Virginia & Stephen Matthews
2007 *The bilingual child: Early development and language contact.* Cambridge: Cambridge University Press.

Guro Nore Fløgstad, Drammen (Norway)
Elizabeth Lanza, Oslo (Norway)

16. Uninstructed language acquisition in multiple language learners

1. What is language acquisition, and why is it relevant to contact linguistics?
2. Methodological preliminaries
3. Acquiring language varieties throughout the lifespan
4. Conclusion
5. References

1. What is language acquisition, and why is it relevant to contact linguistics?

Language acquisition refers to the process whereby an individual learns to understand, speak, interpret signs, sign, read, and/or write. The object of learning may be a variety of a spoken, signed, or written language. This chapter concerns the process of acquisition of multiple oral language varieties (henceforth: languages; see Chevrot and Ghimenton [2019] for a discussion of the extent to which acquisition studies need to distinguish between different ways of speaking).

The chapter focuses on language acquisition in 'naturalistic' settings where individuals need to develop proficiency in a language for interpersonal communication. Such settings are not primarily aimed at systematically helping learners to acquire proficiency in a language (for language acquisition in instructed settings, see Loewen 2015). The acquisition of multiple languages is the locus of language contact at the individual level.

Language acquisition crucially involves change over time towards a state of language knowledge and use that is more extended, and thus more developed, than a previous state. This is expressed in 'language development', a term usually reserved for children's language learning process in early childhood, that is, under the age of 6 (Brooks and Kempe 2014). The more general term 'language learning' stresses the active engagement of learners in the language acquisition process. I use acquisition, development, and learning here as fundamentally synonymous.

There are clear gradations within and amongst learners in the amount of progress they have made towards learning (to use) a particular language. One can thus speak of incipient, intermediate, and advanced learners of a particular language. The rate of language learning, that is, the rate at which people learn new sounds, new words, new structures, new rhetorical devices, new pragmatic functions, etc., generally seems to slow down as a function of the time with which they have engaged with a particular variety of language. However, there is no absolute 'end-state' in language learning (Klein 1986). Thus, language acquisition is never complete. Given intact cognitive skills, good audition, and verbal engagement, all of us continue to learn more language across the lifespan, whether in just one or in more than one. We continue to develop our lexical repertoires (Birdsong 2009). We may create new words that may even be adopted by others. Healthy elderly people tend to be better narrators, to have more expanded semantic networks, and to place more emphasis on meaningful communication of past events than on concise factual representation (Thornton and Light 2006). And as an article in *Forbes*

https://doi.org/10.1515/9783110435351-016

magazine reminds us (Gallo 2014), great orators aren't born as such, but hone their public speaking skills over time through hard work and a lot of practice. Also, the supposed 'effortless' acquisition of language in children under age 6 is not effortless in the least (Clark 2003b; De Houwer 2014). Gaining expertise in speaking takes sufficient learning opportunities and many hours of practice, and using a language equates practicing it (Clark 2003a). As such, all language users, regardless of level of proficiency, are also language learners, and all language learners are language users.

Speaking is not the only language related skill that must be learned. Acquiring a language also involves learning to understand it. Active engagement with others in dyadic interaction provides learning opportunities for both comprehension and production (De Houwer 2009; Ortega 2009).

Developmentally, learning to understand usually comes before learning to speak. Learning to understand language involves learning to perceive speech sounds in meaningful ways. Meaningful perception goes hand in hand with grouping, that is, categorizing, certain speech sounds in the speech signal as belonging together. Such language categorization is already evident in newborns (Brooks and Kempe 2014). Newborn babies who heard either just English or both English and Tagalog prenatally were able to distinguish between English and Tagalog (Byers-Heinlein, Burns, and Werker 2010). However, only the babies with prenatal contact to English and Tagalog showed an equal interest in both languages; the babies with prenatal contact to only English showed a strong preference for English over Tagalog. The prenatally bilingually-exposed babies' interest in both languages from the very start offers an important foundation for their further acquisition of those languages (Byers-Heinlein, Burns, and Werker 2010).

Already at the earliest opportunity for language contact (viz., hearing two languages *in utero*), then, individuals can learn to categorize certain sounds as belonging together but others not, and, what's more, they can attend to each of them equally. Exactly which cues bilingual newborns rely on to separate out their two languages still remains to be investigated. Research with somewhat older bilingually-reared preverbal infants shows that prosodic differences play an important role in helping infants with the fundamental task of learning to distinguish between their languages (Gervain and Werker 2013). For contact linguistics, knowing more about the kinds of cues that bilingual infants use to distinguish between languages may help inform debates about the boundaries between particular languages.

The specific features of languages everywhere are the result of repeated processes of acquisition at the individual level that have been transmitted from generation to generation throughout human history (Brighton, Smith, and Kirby 2005). In learning to speak, people do not simply copy what they hear, but take their input as a basis for constructing their own ways of speaking. After some time, these ways of speaking will show quite a bit of overlap with how others in the same family or community speak, but will also subtly differ, opening the door to language change.

It is only recently that scholars have tried to specifically examine the links between language learning in individuals who are acquiring multiple languages and diachronic language change (Meisel, Elsig, and Rinke 2013). It has long been recognized, though, that some of the acquisition processes taking place in learners of multiple languages contribute to language contact phenomena at the societal level (Treffers-Daller and Mougeon 2005). In particular, repeated patterning of particular types of cross-linguistic influence, that is, the use of linguistic elements from one language when learners are speaking

another, may lead to language change if these individual patterns are shared by many multiple language learners acquiring the same languages. Other contact phenomena in learners' repertoires, such as the use of utterances combining lexical material from two languages, may play a role in language change as well. A typical example is the phenomenon of borrowing at the individual level, which can lead to specific words and phrases from another language being integrated into a 'host' language (Hoffer 2002).

Before I discuss specific aspects of the acquisition of multiple languages, I briefly describe some methodological issues pertaining to language acquisition research. This research is squarely placed in the discipline of psycholinguistics, but has connections with applied linguistics, and to a much lesser degree, with sociolinguistics.

2. Methodological preliminaries

Studying language acquisition involves studying change over time. Thus, the same people should be followed over a longer period of time. The study of language development began with the longitudinal study of individual children. The earliest such study for the development of multiple languages is the one by Ronjat (1913), which is still an inspiring example for scholars today. His study covered his German-French learning son's first two years. Although the bulk of child language acquisition research focuses on the monolingual acquisition of English, it is Leopold's (1939–1949) monumental and insightful longitudinal case studies of German-English bilingual acquisition covering many years that have inspired many influential theories of child language acquisition about how early phonological, lexical, semantic, and mophosyntactic development proceeds.

The close description of language acquisition processes over time is very labor intensive, from data collection through transcribing, coding, and analysis to the actual description. Even though individual longitudinal case studies are still the optimal method for laying bare acquisition processes, which are by nature intensely individual, resources are often lacking to carry them out. There is also the issue of generalizability. With the advent of shared data platforms such as CHILDES (MacWhinney 2000), scholars everywhere have access to longitudinal (and other) naturalistic data that allow investigation of which acquisition processes might be more general, and which are more idiosyncratic.

In order to avoid the need to follow individuals as they grow older (and thus having to wait a long time for results), researchers in language development often carry out cross-sectional studies. These rely on data from different individuals at different stages of life. The assumption is that cross-sectional data represent what happens within the same individual over time. Multiple case studies are a possibility here.

More and more, however, scholars study groups of similarly-aged subjects (regardless of whether they are children or older individuals). Studies may include brief observations of participants interacting with others. More typically, though, aspects of participants' language proficiency and linguistic behavior (such as vocabulary size, the use of specific morphemes, or the use of particular pronunciations) are measured through tests or other standardized instruments, yielding what is termed 'language outcomes'. Typically, these studies compare different groups of learners to each other as a group and do not focus on individual patterns of development over time (De Houwer 2018). Acquisition scholars also use experimental methods to uncover processes that can help us understand aspects of language acquisition.

What is lost in depth for the phenomena that one can realistically study through group studies is gained in potentially improved generalizability. Studies focusing on language outcomes contribute towards an understanding of language acquisition processes, especially if they are longitudinal, but in focusing solely on average group levels much of the individual picture is lost. Group studies have uncovered, however, that similarly-aged language learners differ widely amongst each other in whatever aspect of language use is being measured. Whether this wide inter-individual variation in language outcomes relates to fundamentally different developmental patterns in how languages are learned remains unexplored. Instead, studies focus on factors that are not directly related to language, such as gender or socio-economic status, and on learners' language learning environments, such as the length of overall time that learners have had to learn a language or the specific interactional contexts in which they learn each language (for children, see Armon-Lotem and Meir [2019] and De Houwer [2018]; for adults, see Singleton and Pfenninger [2019]).

In this chapter, I focus mainly on findings from studies that yield insight into the actual trajectories of the acquisition process in individual language learners, rather than on group studies that only measure language outcomes at one specific time.

Depending on one's stage in the lifespan, there are different naturalistic settings in which people come into contact with and can start to learn multiple languages. As my brief and selective review below shows, acquisition-oriented studies of multiple language learners mostly concern children under age 6 or 7, that is, before they have generally started formal literacy instruction. On the whole, we know relatively little about the course of multiple language development in individuals beyond age 6 who are acquiring languages without explicit, formal instruction.

3. Acquiring language varieties throughout the lifespan

3.1. Early childhood: The first 6 years of life

3.1.1. Bilingual First Language Acquisition (BFLA)

As shown above, unborn babies may hear two languages *in utero*. We do not know how common this is. We do know that the setting in which infants are addressed in two languages from birth is a rather common one (De Houwer 2017). This Bilingual First Language Acquisition (BFLA) setting allows children to learn to understand and speak two languages simultaneously from the very beginning (De Houwer 2009). They learn to understand and say words in two languages, and develop two sound systems, thus acquiring two first languages. BFLA children learn to understand and produce translation equivalents from early on. Translation equivalents are words from two languages that mean more or less the same, like French *pomme* and German *Apfel*, meaning 'apple'. BFLA children are able to adjust their language choice according to the communicative setting from around the second birthday (see De Houwer 2009 for an overview). That is, they often will restrict one language Alpha to one particular person, and change to the other language A in addressing another person who speaks language A to them.

Gervain and Werker (2013) showed that already at 7 months, BFLA children are able to distinguish between two fundamentally different word orders. This early ability to notice different morphosyntactic structures may help explain why from the beginning of their own sentence production BFLA children basically treat each of their languages as a morphosyntactically closed set, as proposed in the Separate Development Hypothesis (De Houwer 1990: 66).

3.1.2. Early Second Language Acquisition (ESLA)

Early Second Language Acquisition (ESLA) refers to the setting where children start off hearing and acquiring a single language (L1), and prior to age 6 in addition start to hear a new language, the L2 (De Houwer 1990). Typically, the L2 is used in an institutional setting, such as daycare or preschool (I here use the term L2 for any language learned after another language has already started to be acquired). When ESLA children start to acquire their L2 they are by definition older than they were when they first started to learn their L1. This older age means they are also more socially savvy and more cognitively mature. An effect of the latter is seen in the fact that five-year-old ESLA children make more rapid gains in their L2 vocabulary than monolingual children at ages younger than 5 do in their L1 (Uchikoshi 2006).

The little developmentally oriented research available on ESLA children under age 6 suggests the following developmental path for learning their L2 (what happens to their L1 is even less well known): ESLA children may start off by just learning to understand what is going on in the L2 before attempting to speak it (Gibbons 1985). When they do start speaking, they often use set formulae and single words (Meisel 2008). Gradually, they start to produce sentences that steadily increase in length and complexity. The rate and speed with which ESLA children develop fluency and communicatively adequate skills in the L2 is highly variable (De Houwer 2017).

ESLA children will have formed perceptual categorizations based on their input in just a single language before coming into regular contact with a second one. Depending on their rate of L1 development and the time of first regular contact with an L2, ESLA children may have learned to understand and say some things in their L1 before they start to learn the second. Whatever their level of L1 proficiency, in acquiring their L2 ESLA children build on what they already know in their L1, showing cross-linguistic influence: A study of 1,062 bilingual preschoolers showed that their overall L1 (Spanish) proficiency predicted their L2 (English) proficiency a year later (Winsler, Kim, and Richard 2014).

Another example of cross-linguistic influence may be seen in the sentences that ESLA children produce once they have passed the stage where they produce formulae and single words. Many of those L2 sentences contain clear and systematic influence from morphosyntactic structures belonging to children's L1 (Li Wei 2011; Schwartz and Rovner 2015). Phonological cross-linguistic influence is also common (Hambly et al. 2013). These forms showing cross-linguistic influence may persist as children grow older, or may be replaced by the 'correct' forms as children gain expertise in their L2.

3.1.3. Comparing BFLA and ESLA in preschoolers

Survey studies done in Western Europe and the United States suggest that BFLA occurs three times as often as ESLA (De Houwer 2017). We also know a lot more about BFLA than ESLA children in the first 6 years of life. Nevertheless, it is already possible to formulate some general differences and similarities amongst them.

A notable difference between preschool aged BFLA and ESLA children is that the latter usually do not babble in their L2. For BFLA children, babbling is typically part and parcel of their dual language development (De Houwer 2009). Babbling is typically seen as a necessary stage in children's language development towards the end of the first year of life, but there is debate as to why it is necessary. Its general absence in children beyond the second birthday who start to learn a new language from scratch supports interpretations of babbling as articulatory practicing rather than a necessary step in the acquisition of particular phonological patterns.

Another difference relates to BFLA and ESLA children's use of morphosyntax. The kinds of morphosyntactic forms in the L2 showing cross-linguistic influence from the L1 that young ESLA children produce are largely absent from young BFLA children's speech. In the first years of L2 learning, most of young ESLA children's L2 sentences do not sound like those of monolingual children acquiring that language as an L1. In contrast, the morphosyntactic features of BFLA children's utterances with lexical material from only a single language are generally highly similar to those that monolingual children produce in each of their languages.

There are also basic similarities between BFLA and ESLA preschoolers. Both BFLA and ESLA children use mixed utterances, which combine lexical material from both languages. The proportion of mixed utterances is highly variable, with some children producing hardly any, and others a lot. The reasons for these differences are likely related to language socialization patterns and vocabulary gaps (De Houwer 2009). The majority of mixed utterances in early childhood consists of insertions of a single word from language X into an utterance that is otherwise in language Y, and more often than not, that single word is a noun.

Once they are able to produce short sentences in two languages, both BFLA and ESLA children select the socially appropriate language to speak to specific people. Depending on how children are socialized, they may speak two languages to a specific person. The youngest of bilinguals hardly ever speak a language that their addressee does not understand.

One may ask whether by the time they are old enough to start primary school BFLA and ESLA children have reached similar levels of development in each of their languages. Generally, the answer is no. A major reason is that there is great inter-individual variability amongst children, both within BFLA and within ESLA. Furthermore, there is a lot of intra-individual variability as well. BFLA children's two languages do not necessarily develop at the same rate. Instead, uneven development is to be expected, that is, one language is often better developed than the other (De Houwer 2009). The relation may quickly change, for instance, in response to trips. For instance, a BFLA girl who spoke fluent English but very little Bulgarian rapidly gained in Bulgarian speaking fluency during a trip to Bulgaria but no longer used any English (Slavkov 2015). BFLA children's uneven development may be a lifelong feature, but many BFLA children develop high levels of proficiency in both their languages. ESLA children definitely

experience uneven development at the beginning: Their L1 is much stronger than their incipient L2.

There is more and more research on both BFLA and ESLA in early childhood. Useful reviews are Hammer et al. (2014) and Serratrice (2019). There has also been increasing research attention to children under age 6 who grow up with three languages (Chevalier 2015). The most common pattern there appears to be one where children hear two languages from birth at home, and a third language is introduced some time afterwards through a childminder, daycare, or preschool. Thus, there is often a combination of BFLA and ESLA. A very specific case of ESLA is that of international adoptees, whose learning of an L2 typically coincides with the cessation of exposure to their L1 (Genesee and Delcenserie 2016). International adoptees may be very young when they are adopted, or already of school age. The next section looks at multiple language development in school-aged children, that is, in children aged 6 or 7 and older.

3.2. Multiple language development in school-aged children

Many children arrive in school already knowing two or more languages. These children may be BFLA or ESLA children, as discussed in Section 3.1.

Relatively little research has been done that specifically targets BFLA children beyond the age of 6. So far, there is no evidence of systematic and/or widespread cross-linguistic influence in older BFLA children's use of morphosyntax (none is evident in studies following the same BFLA children from early childhood through adolescence, e.g., Caldas 2006; see also De Houwer 2009). A group study looking at outcomes for vocabulary, on the other hand, found that the more words primary school-aged BFLA children understood in Spanish, the more words they understood in English (Umbel and Oller 1994). This shows that different domains of language use may show different kinds of cross-linguistic relations. Based on cross-sectional evidence, Umbel and Oller (1994) also found that BFLA children's lexical comprehension abilities in their weaker language can become much stronger over the primary school years, though the growth of lexical comprehension abilities in a weaker language need not happen at the expense of the previously stronger language (see Oller, Pearson, and Cobo-Lewis [2007], whose cross-sectional bilingual primary school sample included both BFLA and ESLA children). As ESLA children grow older and start to gain more and more expertise in their L2, both languages may become equally strong, but often, as in BFLA, one language is better developed, at least in production (Oller, Pearson, and Cobo-Lewis 2007). Chesterfield and Chesterfield (1985) offer a rare observational study of ESLA children in preschool with a follow-up when these children were in first grade, and show how the same children used an ever-expanding range of pragmatic strategies across time that went hand in hand with their developing L2 proficiency. There is, however, comparatively little research focusing on the developmental trajectories of oral language use in primary school-aged BFLA and/or ESLA children. We do know that in primary school-aged BFLA and ESLA children the stronger language tends to be the language of schooling. In fact, both BFLA and ESLA school-aged children of any age may end up speaking just a single language. That single language is always the school language. Furthermore, survey data suggest that it is far more common for BFLA than for ESLA children to speak just the school language: A recalculation of the data in De Houwer's (2007: 419)

Table 7 shows that 30 % of children who heard their parents speak both the school language and another language at home (normally, BFLA children) only spoke the school language, whereas only 3 % of children growing up in families where the parents spoke another language than the school language at home (ESLA children) only spoke the school language (see also De Houwer 2017).

Monolingual children above the age of 6 who move to another region or country often meet up with a new language (L2) at school. So do children who have not moved, but who did not have any access to the L2 earlier (e.g., French-speaking families in officially bilingual Brussels may send their monolingual French-speaking children to Dutch-medium schools). In many schools, literacy instruction starts at the same time that children start to learn their new language. Sometimes there are L2 support programs for children who do not yet know the school language. Often, though, it's 'sink or swim' as regards learning to understand the L2 and learning to speak it. In this Second Language Acquisition (SLA) setting, children are surrounded by the L2 at school and have to learn it through listening and trying to interact in it, without explicit instruction or guidance. In addition to hearing the L2 at school, there may be opportunities to hear it and interact in it outside school.

Many studies incorporating measures of oral language use in bilingual school populations do not furnish any language learning history information and focus only on the L2. It is often impossible to know whether bilingual participants learned the L2 from birth, from early childhood, or only after they entered school. This is one reason why we know little about how children acquire oral proficiency in a new language beyond the age of 6 − and even less about how their L1 or any additional languages evolve through the school years. Furthermore, there are very few studies documenting the specific linguistic features of L2 development in school children (Saunders and O'Brien 2006). Thus, for instance, we do not know to what extent SLA students use the kinds of formulaic expressions that appear to be common in ESLA children when they first start to speak the L2 (Myles, Mitchell and Hooper [1999] show that such formulae appear in 11-year-olds learning a foreign language through instruction, though). We do know there is wide variability amongst SLA primary school students in how fast they develop from having no L2 speaking proficiency to having proficiency on a par with monolingual peers: In a rare longitudinal study, MacSwan and Pray (2005) found that children can take as little as one year or as many as 6.5 years to achieve high levels of proficiency in the L2, with 92 % of children taking 5 years. The older children are when they start acquiring the L2, the less time they need to achieve monolingual-like proficiency (MacSwan and Pray 2005). Thus, the commonly heard phrase 'the earlier, the better' finds no empirical support.

3.3. Second Language Acquisition beyond the school years

Many people learn a foreign language at school. People may later emigrate to a country that has that foreign language (FL) as a majority language (the L2). Large scale Canadian and European studies show that adult immigrants with pre-migration FL capital develop higher levels of L2 proficiency much sooner than immigrants without pre-migration FL capital ('pre-migration FL capital' refers to individuals already having learned the new L2 as an FL before migration; Adamuti-Trache 2013; Kristen, Mühlau, and Schacht

2016). Longitudinal case studies of adults with pre-migration FL capital show that they can develop very advanced levels in that same language after migration (for a review, see Ortega 2009).

Newly arrived adult immigrants will often need to learn a fully unfamiliar language in their new country. In some cases, instructional provisions will be in place to help them in the initial stages of L2 language development (Simpson 2019). Very often, though, immigrants are left to their own devices, and learn to understand and speak the L2 through naturalistic interaction and media (Gonçalves 2019; Singleton and Pfenninger 2019).

Newly arrived immigrant adults with different language backgrounds who received no instruction in the L2 may develop a communicatively adequate but morphosyntactically rudimentary L2 system (termed the 'Basic Variety') after a year and a half of exposure (Perdue 1993). Meisel, Clahsen, and Pienemann (1981) found considerable variation amongst the kinds of structures that immigrant adults used. Two thirds of the 40 adults studied in Perdue (1993) moved beyond the Basic Variety, whereas a third did not. Much larger scale longitudinal studies (N = 6,090, Adamuti-Trache [2013]; N = 2,261, Kristen, Mühlau, and Schacht [2016]) confirm the wide variability amongst adult learners in the speed with which they gain a high(er) level of proficiency in the L2. After four years, 47 % of immigrants with low L2 speaking skills on arrival had high L2 speaking skills (immigrants had arrived in Canada at 20 to 59 years of age; Adamuti-Trache 2013). However, 53 % of immigrants with low L2 speaking skills on arrival still had low speaking skills four years later.

Adamuti-Trache (2013) showed that gains in L2 speaking skills were greatest in the first two years after arrival, and levelled off afterwards. In their cross-sectional study of 801 adult immigrants in the United States who had come from Mexico and Cuba, Bahrick et al. (1994) also found that the largest improvements in English L2 proficiency happened in the first years after immigration, although the immigrants continued to improve for many years after (the sample included people up to 70 years of age). Bahrick et al. (1994) also examined Spanish L1 proficiency, and on the whole found no declines as a function of length of residence in the United States. Like for BFLA and ESLA, on the individual level participants' two languages showed uneven development, with many people doing better on L2 than on L1 proficiency measures. Improvements in L2 proficiency were not at the cost of L1 proficiency, however. This reflects findings about children in early childhood (Section 3.1).

4. Conclusion

Learning to speak any language is a long, drawn-out process. For instance, a five-year-old (monolingual) child needs five full years to learn to speak like a five-year-old, with all the limitations that entails (De Houwer 2014). No child beyond preschool age, no adolescent and no adult wants to sound like a five-year-old, and there is no reason they should. After only 5 years of part-time exposure to an L2 through primary school, many children achieve a level of L2 proficiency that is similar to that of 11−12-year-olds who have that language as their only language (Section 3.2), thus needing less than half the time. And many adult immigrants achieve high speaking ability in a new L2 after only

4 years of part-time exposure (Section 3.3), needing even far less time. Generally, then, there is a clear acceleration in new language learning with increasing age. However, regardless of age level, there is a lot of inter-individual variability in the rate of L2 learning and in the specific forms that learners produce. Scholars are trying to explain this variability in terms of differences in, for instance, language exposure (Uchikoshi 2006), social networks (Kristen, Mühlau, and Schacht 2016), media use (Adamuti-Trache 2013), level of proficiency in the other language (Umbel and Oller 1994), the kind of language input that learners hear (De Houwer 2009), and learner-internal factors such as socio-emotional and behavioral skill (Winsler, Kim, and Richard 2014) and gender (Adamuti-Trache 2013; Kristen, Mühlau, and Schacht 2016). Theoretical approaches to the acquisition of multiple languages need to take factors like these into account (e.g., Leather and van Dam 2003). Combined with many more longitudinal studies documenting individuals' language learning trajectories, only approaches that fully take into account the actual context in which people live and their individual characteristics will provide us with more insight about how people learn to use new languages and develop their existing proficiency.

Acknowledgments

I thank Lourdes Ortega for help in finding some of the sources and I am grateful to two anonymous reviewers, whose comments helped improve the manuscript.

5. References

Adamuti-Trache, Maria
2013 Language acquisition among adult immigrants in Canada: The effect of pre-migration language capital. *Adult Education Quarterly* 63. 103–126.

Armon-Lotem, Sharon & Natalia Meir
2019 The nature of exposure and input in early bilingualism. In Annick De Houwer & Lourdes Ortega (eds.), *The Cambridge handbook of bilingualism*, 193–212. Cambridge: Cambridge University Press.

Bahrick, Harry, Lynda Hall, Judith Goggin, Lorraine Bahrick & Stephanie Berger
1994 Fifty years of language maintenance and language dominance in bilingual Hispanic immigrants. *Journal of Experimental Psychology: General* 123(3). 264–283.

Birdsong, David
2009 Age and the end state of second language acquisition. In William Ritchie & Tej Bhatia (eds.), *The new handbook of second language acquisition*, 401–424. Bingley: Emerald Group Publishing.

Brighton, Henry, Kenny Smith & Simon Kirby
2005 Language as an evolutionary system. *Physics of Life Reviews* 2. 177–226.

Brooks, Patricia & Vera Kempe (eds.)
2014 *Encyclopedia of language development*. Thousand Oaks: Sage.

Byers-Heinlein, Krista, Tracey Burns & Janet Werker
2010 The roots of bilingualism in newborns. *Psychological Science* 21(3). 343–348.

Caldas, Stephen
2006 *Raising bilingual-biliterate children in monolingual cultures*. Clevedon: Multilingual Matters.

Chesterfield, Ray & Kathleen Chesterfield
1985 Natural order in children's use of second language learning strategies. *Applied Linguistics* 6(1). 45–59.
Chevalier, Sarah
2015 *Trilingual language acquisition: Contextual factors influencing active trilingualism in early childhood.* Amsterdam: John Benjamins.
Chevrot, Jean-Pierre & Anna Ghimenton
2019 Bilingualism and bidialectalism. In Annick De Houwer & Lourdes Ortega (eds.), *The Cambridge handbook of bilingualism*, 510–523. Cambridge: Cambridge University Press.
Clark, Eve
2003a Critical periods, time, and practice. *University of Pennsylvania Working Papers in Linguistics* 9(2). 39–48.
Clark, Eve
2003b *First language acquisition*. Cambridge: Cambridge University Press.
De Houwer, Annick
1990 *The acquisition of two languages from birth: A case study.* Cambridge: Cambridge University Press.
De Houwer, Annick
2007 Parental language input patterns and children's bilingual use. *Applied Psycholinguistics* 28(3). 411–424.
De Houwer, Annick
2009 *Bilingual first language acquisition*. Bristol: Multilingual Matters.
De Houwer, Annick
2014 Early foreign language teaching: Some critical remarks and some recommendations. *Babylonia* 01/14. 14–21.
De Houwer, Annick
2017 Input, context and early child bilingualism: Implications for clinical practice. In Amalia Bar-On & Dorit Ravid (eds.), *Handbook of communication disorders: Theoretical, empirical, and applied linguistic perspectives*, 599–616. Berlin: Mouton de Gruyter.
De Houwer, Annick
2018 The role of language input environments for language outcomes and language acquisition in young bilingual children. In David Miller, Fatih Bayram, Jason Rothman & Ludovica Serratrice (eds.), *Bilingual cognition and language: The state of the science across its subfields*, 127–153. Amsterdam: John Benjamins.
Gallo, Carmine
2014 Steve Jobs and Winston Churchill didn't start out as great speakers, *Forbes*, Nov. 25, 2014 (http://www.forbes.com/sites/carminegallo/2014/11/25/steve-jobs-and-winston-churchill-didnt-start-out-as-great-speakers/#470cf8f2580a, last accessed March 3, 2018)
Genesee, Fred & Audrey Delcenserie (eds.)
2016 *Starting over: The language development in internationally-adopted children.* Amsterdam: John Benjamins.
Gervain, Judith & Janet Werker
2013 Prosody cues word order in 7-month-old bilingual infants. *Nature Communications* 4(1490). DOI: 10.1038/ncomms2430
Gibbons, John
1985 The silent period: An examination. *Language Learning* 35. 255–267.
Gonçalves, Kellie
2019 Young bilingual adults. In Annick De Houwer & Lourdes Ortega (eds.), *The Cambridge handbook of bilingualism*, 59–75. Cambridge: Cambridge University Press.

Hambly, Helen, Yvonne Wren, Sharynne McLeod & Sue Roulstone
2013 The influence of bilingualism on speech production: A systematic review. *International Journal of Language and Communication Disorders* 48. 1–24.
Hammer, Carol, Erika Hoff, Yuuko Uchikoshi, Cristina Gillanders, Dina Castro & Lia Sandilos
2014 The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly* 29. 715–733.
Hoffer, Bates
2002 Language borrowing and language diffusion: An overview. *Intercultural Communication Studies* XI(4). 1–37.
Klein, Wolfgang
1986 *Second language acquisition*. Cambridge: Cambridge University Press.
Kristen, Cornelia, Peter Mühlau & Diana Schacht
2016 Language acquisition of recently arrived immigrants in England, Germany, Ireland, and the Netherlands. *Ethnicities* 16(2). 180–212.
Leather, Jonathan & Jet van Dam (eds.)
2003 *Ecology of language acquisition*. Dordrecht: Kluwer.
Leopold, Werner
1939–1949 *Speech development of a bilingual child: A linguist's record* (4 volumes). Evanston: Northwestern University Press.
Li Wei
2011 The early acquisition of English as a second language: The case of young Chinese learners of English in Britain. In Annick De Houwer & Antje Wilton (eds.), *English in Europe today: Educational and sociocultural perspectives*, 105–122. Amsterdam: John Benjamins.
Loewen, Shawn
2015 *Introduction to instructed second language acquisition*. New York: Routledge.
MacSwan, Jeff & Lisa Pray
2005 Learning English bilingually: Age of onset of exposure and rate of acquisition among English language learners in a bilingual education program. *Bilingual Research Journal* 29(3). 653–678.
MacWhinney, Brian
2000 *The CHILDES project. Tools for analyzing talk*, 3rd edn. Mahwah: Lawrence Erlbaum Associates.
Meisel, Jürgen
2008 Child second language acquisition or successive first language acquisition? In Belma Haznedar & Elena Gavruseva (eds.), *Current trends in child second language acquisition: A generative perspective*, 55–80. Amsterdam: John Benjamins.
Meisel, Jürgen, Harald Clahsen & Manfred Pienemann
1981 On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition* 3(2). 109–135.
Meisel, Jürgen, Martin Elsig & Esther Rinke
2013 *Language acquisition and change: A morphosyntactic perspective*. Edinburgh: Edinburgh University Press.
Myles, Florence, Rosamund Mitchell & Janet Hooper
1999 Interrogative chunks in French L2: A basis for creative construction? *Studies in Second Language Acquisition* 21. 49–80.
Oller, D. Kim, Barbara Pearson & Alan Cobo-Lewis
2007 Profile effects in early bilingual language and literacy. *Applied Psycholinguistics* 28(02). 191–230.
Ortega, Lourdes
2009 *Understanding second language acquisition*. Abingdon: Routledge.

Perdue, Clive (ed.)
1993 *Adult language acquisition: Cross-linguistic perspectives. Volume 1, Field methods*. Cambridge: Cambridge University Press.
Ronjat, Jules
1913 *Le développement du langage observé chez un enfant bilingue*. Paris: Champion.
Saunders, William & Gisela O'Brien
2006 Oral language. In Fred Genesee, Kathryn Lindholm-Leary, William Saunders & Donna Christian (eds.), *Educating English language learners. A synthesis of research evidence*, 14–63. Cambridge: Cambridge University Press.
Schwartz, Mila & Hadas Rovner
2015 The acquisition of definiteness in Hebrew (L2) by bilingual preschool children (Russian-L1): A longitudinal multiple-case study. *International Journal of Bilingualism* 19(5). 548–571.
Serratrice, Ludovica
2019 Becoming bilingual in early childhood. In Annick De Houwer & Lourdes Ortega (eds.), *The Cambridge handbook of bilingualism*, 15–35. Cambridge: Cambridge University Press.
Simpson, James
2019 Supporting bilingualism in adult first-generation migrants. In Annick De Houwer & Lourdes Ortega (eds.), *The Cambridge handbook of bilingualism*, 250–266. Cambridge: Cambridge University Press.
Singleton, David & Simone Pfenninger
2019 Bilingualism in midlife. In Annick De Houwer & Lourdes Ortega (eds.), *The Cambridge handbook of bilingualism*, 76–100. Cambridge: Cambridge University Press.
Slavkov, Nikolai
2015 Language attrition and reactivation in the context of bilingual first language acquisition. *International Journal of Bilingual Education and Bilingualism* 18. 715–734.
Thornton, Robert & Leah Light
2006 Language comprehension and production in normal aging. In James E. Birren & K. Warner Schaie (eds.), *Handbook of the psychology of aging, sixth edition*, 261–287. New York: Academic Press.
Treffers-Daller, Jeanine & Raymond Mougeon
2005 The role of transfer in language variation and change: Evidence from contact varieties of French. *Bilingualism: Language and Cognition* 8(02). 93–98.
Uchikoshi, Yuuko
2006 English vocabulary development in bilingual kindergarteners: What are the best predictors? *Bilingualism: Language and Cognition* 9(01). 33–49.
Umbel, Vivian & D. Kim Oller
1994 Developmental changes in receptive vocabulary in Hispanic bilingual school children. *Language Learning* 44. 221–242.
Winsler, Adam, Yoon Kim & Erin Richard
2014 Socio-emotional skills, behavior problems, and Spanish competence predict the acquisition of English among English language learners in poverty. *Developmental Psychology* 50. 2242–2254.

Annick De Houwer, Erfurt (Germany)

17. First language attrition and contact linguistics

1. Introduction
2. Effects of attrition on the native language
3. External predictors of attrition
4. Conclusion
5. References

1. Introduction

When speakers of different languages come into contact, their languages typically begin to interact, leading to phenomena such as code-switching and -mixing, various kinds of lexical, structural, and phonological changes, and in extreme cases, the total loss of a language or creation of an entirely new one (Muysken 1995). The addition of another language into the mind leads to the creation of a unified super-system, where the native language is affected by the presence of a second language (L2) (Cook 2012). As Backus (2014) has stated, once the change in the language has set in, it is easy to spot, but its actual origins and how it has taken hold may remain obscure, particularly given that documentation of earlier, pre-contact stages of the language in question is very often scarce. In addition, the focus of contact-linguistic research is usually at the level of the speech community and macro-societal factors, with a primary interest in either the replacement of the community language by another, or in substantial changes in the linguistic systems of these communities (Lim and Ansaldo 2015). Only rarely do investigations attempt to reveal the complex interactions of linguistic, social, and attitudinal factors, both at the individual and the societal level (Winford 2003), or investigate both the social as well as the cognitive aspects of phenomena, such as code-switching in bilingual individuals (Myers-Scotton 2002).

Consequently, there has been more emphasis on linguistic outcomes at the societal level, and less on how individual speakers' linguistic abilities and practices have developed and changed over time, but as such individual and microlinguistic processes of change constitute the seed of wholesale change and adaptation of linguistic systems, they form an important part of the overall process. For instance, when a speaker starts having difficulties with the grammatical principles of gender concord, this will have wider consequences, because gender involves the interaction of several components (i.e. morphology, syntax, semantics, and phonology) as well as knowledge about the real world (Corbett 1991). When gender and the related features cannot be inferred from the input, this could lead to incorrect usage (e.g. overgeneralization of masculine forms and indefinite articles as in heritage Norwegian, reported by Lohndal and Westergaard [2016]), and eventually, erosion of the gender system and deterioration of related features in the language variety. The aim of this chapter is to point to the role that individual speakers play in situations where language change is taking hold, and which may (or may not) eventually lead to large-scale processes of change and/or loss at the societal level and over generations (Matras 2009).

The phenomenon of first language attrition in the individual was first recognized in the classical work of Haugen (1938: 1) as the possible outcome of a "tug of war" between the languages of the immigrant who arrives in a new language environment. Similarly, Weinreich (1953) emphasized the importance of including the study of individ-

https://doi.org/10.1515/9783110435351-017

ual bilinguals in order to understand the mechanisms of societal language contact and change. However, it took three more decades until language attrition was constituted as a research field. This is generally assumed to have happened with a conference on *Language Loss*, which took place at UPenn in 1980 (Lambert and Freed 1982). Since then, various symposia and conferences, graduate workshops, collected volumes, and special issues in journals have contributed to the visibility of attrition research as an increasingly important subfield of bilingual development (see Köpke and Schmid [2004] for a historical review), and investigations into first language attrition have been recognized as an integral part of language acquisition and bilingualism research.

The term 'attrition', first used by Haugen (1938: 1) for the "slow and incessant" process by which "each foreigner has been turned into an American, idea by idea, and word by word". It describes "any of the phenomena that arise in the native language of a sequential bilingual as the consequence of the co-activation of languages, cross-linguistic transfer or disuse, at any stage of L2 development and use" (Schmid and Köpke 2017: 637). Unfortunately, its strong and largely negative connotations appear to have led to the widespread perception that attrition is a comparatively rare process, confined to extreme and prolonged situations of very limited L1 use (Costa and Sebastián-Gallés 2014) coupled with very high levels of L2 proficiency (Kroll and Bialystok 2013). In other words, it is often assumed that 'actual' attrition should somehow go beyond the "normal influence between languages in a bilingual or polyglot" (Ahlsén 2013: 1). We hold that such a division of bilinguals into 'non-attriters' and 'attriters' is both impractical and unhelpful: all bilingual speakers experience bi-directional transfer to some extent. Such phenomena can be seen across all linguistic levels, and are an inevitable outcome of the fact that the management of a bilingual system requires more cognitive resources than is the case for monolinguals because both languages are active (to various degrees) at the same time, regardless of the speaker's intention to use one and inhibit the other. Therefore, bilinguals need more time and effort to resolve cross-linguistic influences (Green and Abutalebi 2013), and this sometimes leads to language transfer phenomena, all of which can be considered instances of attrition.

That notwithstanding, the context in which language attrition has most frequently been studied, and which is most relevant to the topic of language contact, is that of immigrants who have migrated to another country and learned the language of the majority society as adults. While they often continue to use both languages in their daily life, the contexts and frequency of L1 use are considerably reduced in comparison to their previous language environments. Most of these speakers have been observed to preserve their general proficiency in the L1 at high levels; yet, they indicated some degree of vulnerability in their linguistic system due to reduction in exposure and use (Schmid 2016). The following sections will provide an overview of the extent to which a speaker's native language can be compromised, and what factors play a role in the process of attrition.

2. Effects of attrition on the native language

There is a considerable body of research demonstrating that attrition phenomena can occur across all linguistic levels (lexicon: Pavlenko 2009; Schmid and Jarvis 2014; morphosyntax: Gürel 2007; Yılmaz 2011; phonetics: Bergmann et al. 2016; de Leeuw, Tusha,

and Schmid 2018; pragmatics: Chamorro, Sorace, and Sturt 2016; Dewaele 2004). It can manifest itself as a simplification or reduction of the L1 system (Gürel 2013), or disfluent speech with a high proportion of pauses and self-repairs (Bergmann, Sprenger, and Schmid 2015). Attrition effects, in the form of structural changes and restructuring, have been reported to be most dramatic for early bilinguals who were exposed to an L2 in childhood or before puberty, see Bylund (2009); Montrul (2016). Among mature speakers, however, attrition outcomes do not usually lead to a wholesale deterioration or loss of a particular linguistic feature or skill. Rather, the speakers seem to have retrieval and processing difficulties due to the cognitive load of bilingualism.

The lexicon is the place where change is the most immediately visible to researchers and bilingual speakers themselves. Bilinguals often have the tendency to borrow words/expressions from the L2, to integrate them phonologically and/or morphologically, to use L2-like collocations and idioms, to extend the meaning of an L1 word to capture the meaning of its L2 translation equivalent, and to converge the L1 term with its L2 meaning (Pavlenko 2004). Such online mergers, however, do not necessarily indicate that the speaker no longer has these words and phrases in the L1 system. Instead, they can be ascribed to effects of cross-linguistic competition during online speech, with both languages contributing to some extent to the process of lexical retrieval or production. Analyses of free speech data further reveal that this competition can lead to difficulties in lexical retrieval, manifesting themselves in reduced levels of lexical diversity (Schmid and Jarvis 2014) and decreased creativity and fluency in speech (Bergmann, Sprenger, and Schmid 2015) compared to monolingual speakers. However, bilinguals also tend to develop compensatory strategies which prevent these processes of transfer and interference from disrupting communication. In controlled experiments, such as picture naming and verbal fluency, where speakers can focus entirely on lexical retrieval, they typically have fewer problems of accessibility (Yılmaz and Schmid 2012). All this implies that there is no dramatic loss of vocabulary (Köpke and Nespoulous 2001; Schmid and Jarvis 2014): the underlying knowledge of the lexical items is intact, but retrieval can sometimes fail in real-time language processing and production due to the cognitive demands of managing both linguistic systems in the face of reduced automaticity.

With respect to grammatical change, similarly, the driving force appears to be a simplification of costly syntactic operations (Gürel 2013; Seliger and Vago 1991; Sorace and Serratrice 2009). Among the phenomena which have most often been studied in this respect are: case morphology (Larmouth 1974; Pavlenko 2003), gender-marking and adjective/noun convergence (Bergmann et al. 2015), determiners (Ben-Rafael 2004), use of relative clauses and complex embeddings (Yılmaz 2011), relative clause attachments (Kasparian and Steinhauer 2017), simplification of word order (Backus and Onar Valk 2013), reduction of restrictions in the binding properties of null versus overt pronouns (Gürel 2007), and elimination of the perfective/imperfective aspectual distinction (Pavlenko 2003). Similar to the findings on the lexicon reported above, most of these studies agree that attrition phenomena are more likely to be an outcome of difficulties with the integration and coordination of information at various levels, rather than an erosion of rules (within the generative framework, this is often put in terms of interface versus core syntax phenomena [Sorace 2011]). In other words, morphosyntactic phenomena are most often ascribed to a reorganization of the linguistic system as a result of integrating the newly acquired L2 structures, allowing bilinguals to efficiently cope with cross-language competition, which is a natural consequence of bilingualism (Chamorro, Sorace, and Sturt 2016).

Similar processes of adaptation can be observed in the area of phonology and in particular, phonetics. A number of investigations of bi-directional influence in speech production, beginning with Flege's (1987) investigation of late French-English and English-French bilinguals, have found evidence for such adaptation. Flege's original findings that Voice-Onset Time in both of a bilingual's languages will shift towards the setting of the other contact language over time and with an increase in proficiency, have since been replicated in studies of Brazilian Portuguese and English (Major 1992), English and Dutch (Schmid, Gilbers, and Nota 2014), and English and Korean (Chang 2012). Bi-directional changes have also been found for the production of some vowels (Bergmann et al. 2016), the production of lateral /l/ in German-English bilinguals (de Leeuw, Mennen, and Scobbie 2013), rhoticity (Himmel and Kabak 2016; Ulbrich and Ordin 2014), and suprasegmentals (Mennen 2004). At the same time, a range of studies on perceived global foreign accent find that attriters are not reliably identified as native speakers by native raters (Bergmann et al. 2016; Hopp and Schmid 2013). Attrition in this domain seems to be more prevalent than other types of attrition, as attriters themselves often report that they are recognized by the way they speak when they return to their country of origin (Yılmaz and Schmid 2012). On the other hand, bilinguals seem to retain the ability to perceive foreign accent in their native language, suggesting that they in fact maintain their underlying phonetic abilities (Major 2010). These findings yet again do not indicate a restructuring of the underlying phonemic space, but suggest that the native speech spectrum has been widened, due to daily exposure to a larger range of different variants of the L1 in an L2 environment (e.g. other attriters, heritage speakers, speakers of other dialects of the L1, and L2 speakers). We do not know of any studies of phonological recovery with re-exposure to target-like forms (as has been found for grammatical phenomena, e.g. Chamorro, Sorace, and Sturt [2016] on subject pronouns), however, it has been proposed that once re-immersed in the L1 environment, speakers would gradually become re-tuned to the non-migrant variant of the L1 (Herdina and Jessner 2013). Whether this is indeed true, and how long this process of re-adaptation would take, is a matter for future research.

In summary, it can thus be said that the process of attrition represents the adaptation of a fully developed, monolingual language system to the demands of competition and limited cognitive resources when a second language is introduced. Such adaptations can be seen across all linguistic levels, affecting production, perception, and comprehension, and represent surface level phenomena. Gradually, these adaptations may lead to shifts in distributional patterns across the speech community, which are then transmitted to subsequent generations and eventually result in an established contact variety. For instance, young speakers of German and Croatian minority dialects in Italy were observed to have irregular patterns of language acquisition and use which was divergent from the traditional one (e.g. overgeneralization of case marker), yet they appeared to have a fully functioning language system. As these speakers extend their domains of use, and as this variety gains respect among the community, it is possible that the (semi)speakers of this language will become fluent speakers of this language variety (Dal Negro 2004). Investigations of early language attrition may thus allow the reconstruction of these early situations of contact-induced language change, or even the pre-contact state of a linguistic system, in cases where documentation of previous stages of a moribund language is lacking.

3. External predictors of attrition

Fishman (1991) points out that the complexity of the process of language change in situations of contact makes it hard to avoid broad overgeneralizations, obscuring the causes and detailed patterns of shift and maintenance. This underscores the importance of investigating the mechanisms of language change and their interactions in more detail. Investigations of language attrition not only find statistically significant differences between attrited and non-attrited populations at group level, but also almost always find a larger range of variability among the individual participants in the bilingual than in the monolingual populations (Schmid 2013). Arguably, such variability constitutes the seed of language change and is the necessary precursor for any processes of shift and adaptation to set in (Croft 2000). Many attrition studies have attempted to account for the degree of individual adaptation to the contact language on the basis of factors related to personal background (age of acquisition, education), to the immigration experience (length of residence, context of migration), to language habits (use and exposure to L1 and L2 in different settings), and to attitudes towards the language and culture of both speech communities.

One factor that is of particular relevance to understanding how individual attrition may lead to contact-induced change in the wider community is language use and exposure. While early studies of attrition more or less axiomatically assumed that higher levels of L1 use would be strongly related with lower degrees of attrition, the impact of amount and frequency of language use in daily life and in domains outside the home has more recently been shown to not play a major role. A set of recent studies, based on a detailed test battery for the measurement of different types of exposure (interactive: spoken and written communication with family, friends, colleagues, and others; noninteractive: reading and media; inner language: thought, dreams, diary writing and arithmetic) yielded few indications of a link between such patterns of exposure and actual language attrition phenomena (Opitz [in press]; for the test battery see languageattrition.org/resources-for-researchers/experiment-materials/sociolinguistic-questionnaires, last accessed on 13 March 2018). Based on these findings, Schmid (2011) suggests that variability in attrition phenomena is related to language use at different degrees of linguistic co-activation, and consequently, more a function of efficient inhibition of the L2 than of failure of activation of the L1. In this model, frequent, informal use of the L1 with other bilingual speakers ('bilingual mode', according to Grosjean's [2001] model of language modes) not only does not protect the language against attrition, but is actually conducive to the development of a contact-influenced variety of the language. When both languages are highly active and code-switching is frequent, those online switches strengthen the associative chains across languages which eventually leads to more intensive interference. Frequent use of the L1 in groups or communities where code-switching is widely practiced may thus lead to an acceleration of contact-induced change − and thus higher levels of individual attrition (Grosjean and Py 1991). On the other hand, speakers who use their L1 in situations where code-switching is discouraged or inappropriate, e.g. in professional settings or language clubs, tend to stick more closely to the target language variety and expend more effort on inhibiting the L2, resulting in lower degrees of co-activation. These types of exposure and use may thus help retention and also foster the development of inhibitory skills (Green 2011; Schmid and Köpke 2017), delaying such changes.

A second set of factors which also plays a very prominent role in language contact settings is socio-psychological in nature and relates to cultural and linguistic identity. It has often been shown that, together with its perceived prestige, the emotional attachment which immigrant groups hold towards their native language and culture, and their willingness to pass it on to future generations, is instrumental in the preservation of a language (Clyne 2003), but it has been extremely difficult to establish similar patterns at the individual level. One of the first studies that did find a clear link between attitudes and attrition was carried out among German Holocaust survivors who migrated to Anglophone countries before the outbreak of World War II (Schmid 2002). This investigation demonstrated that the subset of participants who had experienced the worst levels of oppression and persecution also experienced the highest degrees of L1 attrition. The desire to distance themselves from the L1 community, in this group, outweighed all other variables, including the age of arrival in the host country and the frequency of use of German (see also Ben-Rafael and Schmid 2007). Other studies have found the role of attitudes more difficult to establish (e.g. Hulsen 2000; Opitz 2011; Waas 1996; Yağmur 1997). This divergence in findings may be due to two factors. The first of these is a potential lack of variability within the speech communities under investigation: while Schmid (2002) looks at qualitatively different contexts of migration, and Ben-Rafael and Schmid (2007) investigate two different migrant communities, the other studies cited here look at a single, largely homogenous speech community. This implies that factors such as prestige and institutional language support (which are among the factors analyzed in particular for those studies that are situated within the Ethnolinguistic Vitality framework [Giles, Bourhis, and Taylor 1977]) do not vary within the population, and that it is only the subjective perception of these factors by an individual which can be assessed. The methodological difficulties of such an assessment constitute the second potential problem: studies often rely on questionnaires, and this type of snap-shot measurement may not reliably reflect attitudes which tend to fluctuate across the life span, depending on circumstances. Qualitative analyses of personal histories would provide better insight, as demonstrated by Cherciov (2013). Therefore, it appears that in order to fully explore the link between the attitudes and language profiles of the speakers, wider studies which not only compare different populations and migratory contexts, but also combine quantitative and qualitative approaches, are necessary.

While investigating the role of individual factors, attrition studies also acknowledge that these variables belong to a complex web where they are interrelated with each other and contribute to the overall development of language attrition or retention (Cherciov 2013; Köpke 2007; Opitz in press). Some researchers have recently considered a Dynamic Systems Theory perspective, which takes L2 and L1 as interdependent systems in a continuous process of competition and adaptation in relation to changing circumstances, and aspires to describe the developmental path of a given set of initial linguistic conditions (Cherciov 2013). The case of external variables is inherently vague and imprecise. They do not determine the attritional trajectory in isolation, but continuously fluctuate and interact with each other. Language development is the outcome of interactions between individual factors, where a factor may not necessarily influence language maintenance and shift directly, but only indirectly via intervening variable(s). To the best of our knowledge, no concrete models have been developed within this framework except Meara's (2004) work on an artificial lexical network. Such theoretical models, while less complex than real language, are suggestive of what may happen to an attriter's lexicon.

Therefore, sophisticated simulations that capture more linguistic levels are necessary. These might then be standardized and made cross-linguistically adaptable so that statistical models of language change, and even of reversal (re-learning), could be designed. Such statistical models would help us better understand the mechanisms and nature of language change.

4. Conclusion

The widespread consensus among researchers in the field of contact linguistics that language outcomes are determined largely by the socio-historical relations among populations who are in contact has somewhat overlooked the role of the individual. The latter phenomenon is the starting point of the former: when a community gradually loses its native language and shifts to another language over time, it is in fact the individual speaker whose language ability and dominance first begins to shift (Dorian 1982; Seliger 1996). This is illustrated in Gonzo and Salterelli's (1983) 'cascade' model, where each subsequent generation ends up with a reduced variety of their parents' linguistic repertoire, leading to disappearance of the immigrant language across three or four generations. In order to fully understand both the language-internal and language-external processes which determine the patterns of shift and maintenance, we propose that more attention should be given to the early phases of the process − the point where the original, fully developed monolingual system is first modified and affected by language transfer, and begins to exhibit increased variability.

In this respect, attrition research can play an important role, in particular since it draws on various theoretical and methodological perspectives (e.g. simplification, interlanguage, psycholinguistics, and recently, neurolinguistics) to interpret the data (for overviews see Köpke and Schmid 2004; Schmid 2016). Attrition thus has the potential to provide corroborative evidence for an in-depth, integrated perspective into the process of change: attrition studies rely on a multitude of datasets, ranging from formal tasks that target both production and processing to free-speech analyses that reveal the actual state of the attriting individual, as well as extensive linguistic and personal background interviews (see Schmid [2011] for an overview), and they investigate specific lexical, grammatical, and phonological features as well as general proficiency in a variety of attrition settings. Such systematic investigations are often impossible for the documentation of longer-term historical change, and they help us understand how some (subtle) changes in the native language pave the way for permanent changes/restructuring and establish themselves, while others recede. In this way, attrition research may lead to more principled predictions about language change.

Discovering the underlying principles of L1 change is crucial to making sense of the phenomena witnessed in contact-induced change, because they are similar to L1 attrition (Riehl in press) and they may be the result of the same grammatical principles and processes (Myers-Scotton 2002). Since language change diffuses from individuals to smaller groups and then to the speech community, arriving at a full account of contact phenomena without (re)integrating the individual into the overall matrix of speech community is impossible (Sankoff 2001). Considering them all as part and parcel of the same developmental cycle may better inform our understanding of how language lives, grows, changes, and dies, both in the human mind and in society.

5. References

Ahlsén, Elisabeth
2013 Language attrition. In Carol A. Chappelle (ed.), *The encyclopedia of applied linguistics*, 1–6. Oxford: Blackwell. DOI: 10.1002/9781405198431.wbeal0477.

Backus, Ad
2014 Towards a usage-based account of language change: Implications of contact linguistics for linguistic theory. In Robert Nicolaï (ed.), *Questioning language contact: Limits of contact, contact at its limits*, 91–118. Leiden/Boston: Brill.

Backus, Ad & Pelin Onar Valk
2013 Syntactic change in an immigrant language: From non-finite to finite subordinate clauses in Turkish. *Journal of Estonian and Finno-Ugric Linguistics* 4(2). 7–29.

Ben-Rafael, Miriam
2004 Language contact and attrition: The spoken French of Israeli Francophones. In Monika S. Schmid, Barbara Köpke, Merel C. J. Keijzer & Lina Weilemar (eds.), *First language attrition: Interdisciplinary perspectives on methodological issues*, 165–188. Amsterdam: John Benjamins.

Ben-Rafael, Miriam & Monika S. Schmid
2007 Language attrition and ideology: Two groups of immigrants in Israel. In Barbara Köpke, Monika S. Schmid, Merel C. J. Keijzer & Susan Dostert (eds.), *Language attrition: Theoretical perspectives*, 205–226. Amsterdam: John Benjamins.

Bergmann, Christopher, Amber Nota, Simone A. Sprenger & Monika S. Schmid
2016 L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics* 58. 71–86.

Bergmann, Christopher, Nienke Meulman, Laurie A. Stowe, Simone A. Sprenger & Monika S. Schmid
2015 Prolonged L2 immersion engenders little change in morphosyntactic processing of bilingual natives. *Neuroreport* 26(17). 1065–1070.

Bergmann, Christopher, Simone A. Sprenger & Monika S. Schmid
2015 The impact of language co-activation on L1 and L2 speech fluency. *Acta Psychologica* 161. 25–35.

Bylund, Emanuel
2009 Maturational constraints and first language attrition. *Language Learning* 59(3). 687–715.

Chamorro, Gloria, Antonella Sorace & Patrick Sturt
2016 What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition. *Bilingualism: Language and Cognition* 19(3). 520–532.

Chang, Charles B.
2012 Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of phonetics* 40(2). 249–268.

Cherciov, Mirela
2013 Investigating the impact of attitude on first language attrition and second language acquisition from a Dynamic Systems Theory perspective. *International Journal of Bilingualism* 17(6). 716–733.

Clyne, Michael
2003 *Dynamics of language contact: English and immigrant languages.* Cambridge: Cambridge University Press.

Cook, Vivian J.
2012 Multi-competence. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics*, 3768–3774. Oxford: Blackwell.

Corbett, Greville G.
1991 *Gender.* Cambridge: Cambridge University Press.

Costa, Albert & Núria Sebastián-Gallés
2014 How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience* 15(5). 336–345.
Croft, William
2000 *Explaining language change: An evolutionary approach.* London: Longman.
Dal Negro, Silvia
2004 Language contact and dying languages. *Revue Française de Linguistique Appliquée* 2(10). 47–58.
de Leeuw, Esther, Aurela Tusha & Monika S. Schmid
2018 Individual phonological attrition in Albanian-English late bilinguals. *Bilingualism: Language and Cognition* 21(2). 278–295. DOI:10.1017/S1366728917000025.
de Leeuw, Esther, Ineke Mennen & James M. Scobbie
2013 Dynamic systems, maturational constraints and L1 phonetic attrition. *International Journal of Bilingualism* 17(6). 683–700.
Dewaele, Jean-Marc
2004 The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development* 25(2/3). 204–222.
Dorian, Nancy C.
1982 Defining the speech community in terms of its working margins. In Suzanne Romaine (ed.), *Sociolinguistic variation in speech communities*, 25–33. London: Edward Arnold.
Fishman, Joshua. A
1991 *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon: Multilingual Matters.
Flege, James E.
1987 The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15. 47–65.
Giles, Howard, Richard Y. Bourhis & Donald M. Taylor
1977 Towards a theory of language in ethnic group relations. In Howard Giles (ed.), *Language, ethnicity and intergroup relations*, 307–348. London: Academic Press.
Gonzo, Susan & Maria Saltarelli
1983 Pidginization and linguistic change in emigrant languages. In Roger W. Andersen (ed.), *Pidginization and creolization as language acquisition*, 181–197. Rowley, Massachusetts: Newbury House.
Green, David W.
2011 Language control in different contexts: The behavioral ecology of bilingual speakers. *Frontiers in Psychology* 2. 103. DOI: 10.3389/fpsyg.201.00103.
Green, David W. & Jubin Abutalebi
2013 Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* 25(5). 515–530.
Grosjean, François
2001 The bilingual's language modes. In Janet L. Nicol (ed.), *One mind, two languages: Bilingual language processing*, 1–22. Oxford: Blackwell.
Grosjean, François & Bernard Py
1991 La restructuration d'une première langue: L'intégration de variantes de contact dans la compétence de migrants bilingues. *La Linguistique* 27(2). 35–60.
Gürel, Ayşe
2007 (Psycho)linguistic determinants of L1 attrition. In Barbara Köpke, Monika S. Schmid, Merel C. J. Keijzer & Susan Dostert (eds.), *Language attrition: Theoretical perspectives*, 99–119. Amsterdam: John Benjamins.
Gürel, Ayşe
2013 First language attrition of constraints on wh-scrambling: Does the second language have an effect? *International Journal of Bilingualism* 19(1). 75–91.

Haugen, Einar
1938 Language and immigration. *Norwegian-American Studies and Record* 10. 1–43.
Herdina, Philip & Ulrike Jessner
2013 The implications of language attrition for dynamic systems theory: Next steps and consequences. *International Journal of Bilingualism* 17(6). 752–756.
Himmel, Marie-Cristin & Barış Kabak
2016 Adaptation or attrition? L1 rhoticity in American English-German late bilinguals. Paper presented at the 3rd International Conference on Language Attrition, Colchester, UK, July 2016.
Hopp, Holger & Monika S. Schmid
2013 Perceived foreign accent in L1 attrition and L2 acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics* 34(2). 361–394.
Hulsen, Madeleine E. H.
2000 *Language loss and language processing: Three generations of Dutch migrants in New Zealand*. Nijmegen: Katholieke Universiteit Nijmegen dissertation.
Kasparian, Kristina & Karsten Steinhauer
2017 When the second language takes the lead: Neurocognitive processing changes in the first language of adult attriters. *Frontiers in Psychology* 8. 389.
Köpke, Barbara
2007 Language attrition at the crossroads of brain, mind, and society. In Barbara Köpke, Monika S. Schmid, Merel C. J. Keijzer & Susan Dostert (eds.), *Language attrition: Theoretical perspectives*, 9–37. Amsterdam: John Benjamins.
Köpke, Barbara & Jean-Luc Nespoulous
2001 First language attrition in production skills and metalinguistic abilities in German-English and German-French bilinguals. In Tom Ammerlaan, Madeleine Hulsen, Heleen Strating & Kutlay Yağmur (eds.), *Sociolinguistic and psycholinguistic perspectives on maintenance and loss of minority languages*, 221–234. Münster: Waxmann.
Köpke, Barbara & Monika S. Schmid
2004 Language attrition: The next phase. In Monika S. Schmid, Barbara Köpke, Merel C. J. Keijzer, & Lina Weilemar (eds.), *First language attrition: Interdisciplinary perspectives on methodological issues*, 1–43. Amsterdam: John Benjamins.
Kroll, Judith F. & Ellen Bialystok
2013 Understanding the consequences of bilingualism for language processing and control. *Journal of Cognitive Psychology* 25(5). 497–514.
Lambert, Richard D. & Barbara F. Freed (eds.)
1982 *The loss of language skills.* Rowley: Newbury House.
Larmouth, Donald W.
1974 Differential interference in American Finnish cases. *Language* 50. 356–366.
Lohndal, Terje & Marit Westergaard
2016 Grammatical gender in American Norwegian heritage language: Stability or attrition? *Frontiers in Psychology* 7. 344.
Lim, Lisa & Umberto Ansaldo
2015 *Languages in contact*. Cambridge: Cambridge University Press.
Major, Roy C.
1992 Losing English as a first language. *The Modern Language Journal* 76(2). 190–208.
Major, Roy C.
2010 First language attrition in foreign accent perception. *International Journal of Bilingualism* 14(2). 163–183.
Matras, Yaron
2009 *Language contact.* Cambridge: Cambridge University Press.
Meara, Paul
2004 Modelling vocabulary loss. *Applied Linguistics* 25(2). 137–155.

Mennen, Ineke
2004 Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32(4). 543–563.
Montrul, Silvina
2016 *The acquisition of heritage languages*. Cambridge: Cambridge University Press.
Muysken, Pieter C.
1995 Code-switching and grammatical theory. In Lesley Milroy & Pieter C. Muysken (eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, 177–198. Cambridge: Cambridge University Press.
Myers-Scotton, Carol
2002 *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
Opitz, Cornelia
2011 *First language attrition and second language acquisition in a second-language environment*. Dublin: Trinity College dissertation.
Opitz, Cornelia
In press Personal background factors in L1 attrition: A perspective from Dynamic Systems Theory. In Monika S. Schmid, Barbara Köpke, Mirela Cherciov, Esther de Leeuw, Merel C. J. Keijzer, Teodora Mehotcheva & Tuğba Karayayla (eds.), *The Oxford handbook of language attrition*. Oxford: Oxford University Press.
Pavlenko, Aneta
2003 I feel clumsy speaking Russian: L2 influence on L1 narratives of Russian L2 users of English. In Vivian Cook (ed.), *Effects of the second language on the first*, 32–61. Clevedon: Multilingual Matters.
Pavlenko, Aneta
2004 Second language influence and first language attrition in adult bilingualism. In Monika S. Schmid, Barbara Köpke, Merel C. J. Keijzer & Lina Weilemar (eds.), *First language attrition: Interdisciplinary perspectives on methodological issues*, 47–59. Amsterdam: John Benjamins.
Pavlenko, Aneta (ed.)
2009 The bilingual mental lexicon: Interdisciplinary approaches. Clevedon, UK: Multilingual Matters.
Riehl, Claudia M.
In press Attrition and language contact. In Monika S. Schmid, Barbara Köpke, Mirela Cherciov, Esther de Leeuw, Merel C. J. Keijzer, Teodora Mehotcheva & Tuğba Karayayla (eds.), *The Oxford handbook of language attrition*. Oxford: Oxford University Press.
Sankoff, Gillian
2001 Linguistic outcomes of language contact. In Peter Trudgill, Jack Chambers & Natalie Schilling-Estes (eds.), *Handbook of sociolinguistics*, 638–668. Oxford: Basil Blackwell.
Schmid, Monika S.
2002 *First Language attrition, use and maintenance: The case of German Jews in Anglophone countries*. Amsterdam/Philadelphia: John Benjamins.
Schmid, Monika S.
2011 Contact x time: External factors and variability in L1 attrition. In Monika S. Schmid, & Wander Lowie (eds.), *Modelling bilingualism: From structure to chaos*, 155–176. Amsterdam/Philadelphia: John Benjamins.
Schmid, Monika S.
2013 First language attrition. *WIRE's Cognitive Science* 4(2). 117–123.
Schmid, Monika S.
2016 Language attrition: Timeline. *Language Teaching* 49(2). 186–312.

Schmid, Monika S. & Barbara Köpke
2017 The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism* 7(6). 637–667.
Schmid, Monika S. & Scott Jarvis
2014 Lexical first language attrition. *Bilingualism: Language and Cognition* 17(4). 729–748.
Schmid, Monika S., Steven Gilbers & Amber Nota
2014 Ultimate attainment in late second language acquisition: Phonetic and grammatical challenges in advanced Dutch-English bilingualism. *Second Language Research* 30(2). 129–157.
Seliger, Herbert W.
1996 Primary language forgetting in the context of bilingualism. In William C. Ritchie & Tej K. Bhatia (eds.), *Handbook of second language acquisition*, 605–626. San Diego: Academic Press.
Seliger, Herbert W. & Robert M. Vago
1991 *First language attrition*. Cambridge: Cambridge University Press.
Sorace, Antonella
2011 Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism* 1(1). 1–33.
Sorace, Antonella & Ludovica Serratrice
2009 Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism* 13(2). 195–210.
Ulbrich, Christiane & Mikhail Ordin
2014 Can L2-English influence L1-German? The case of post-vocalic /r/. *Journal of Phonetics* 45. 26–42.
Waas, Margit
1996 *Language attrition downunder*. Frankfurt: Peter Lang.
Weinreich, Uriel
1953 *Languages in contact.* The Hague: Mouton.
Winford, Donald
2003 *An introduction to contact linguistics.* Oxford: Blackwell.
Yağmur, Kutlay
1997 *First language attrition among Turkish speakers in Sydney.* Tilburg: Tilburg University Press.
Yılmaz, Gülsen
2011 Complex embeddings in free speech production among late Turkish-Dutch bilinguals. *Language, Interaction and Acquisition* 2(2). 251–275.
Yılmaz, Gülsen & Monika S. Schmid
2012 L1 accessibility among Turkish-Dutch bilinguals. *Mental Lexicon* 7(3). 249–274.

Gülsen Yılmaz, Berlin (Germany)
Monika S. Schmid, Essex (UK)

18. Individual variation in bilingual lexical processing: the impact of second language proficiency and executive function on cross-language activation

1. Introduction
2. Bilingual lexical processing: theoretical perspectives
3. Bilingual lexical processing: studies manipulating lexical-semantic overlap
4. Cross-language activation and second language proficiency
5. Cross-language activation and executive function
6. Concluding remarks
7. References

1. Introduction

When speakers of different languages interact closely over an extended period of time, their languages influence each other. Language contact can lead to structural changes in these languages (exemplified by language convergence, borrowing, and language shifts), but it can also lead to changes in the speakers' neural and cognitive systems that underlie language processing. This chapter approaches language contact from a psycholinguistic perspective, reviewing selected theoretical models and empirical studies on cross-language interaction of lexical-semantic codes during lexical processing. More specifically, we focus on how the nature of cross-language interactions during lexical processing varies as a function of individual variation in language proficiency and executive function among bilingual speakers. We define bilinguals (or multilinguals) as individuals who regularly use two (or multiple) languages.

2. Bilingual lexical processing: theoretical perspectives

Psycholinguistic research investigating cross-language interaction and lexical processing has long been characterized by testing two competing views on how bilinguals access and activate words in their two languages: language selective and language non-selective views. The language-selective view claims that bilinguals activate only words from the language system that corresponds with the language of the incoming information (in comprehension) or with the language currently in use (in production). In contrast, the non-selective view holds that words from both languages are activated, even when the incoming information or the language in use is restricted to one language only. There is ample evidence that lexical activation in the bilingual mental lexicon operates in a parallel, language non-selective way, even when the social and linguistic context calls for only one language (Van Hell and Tanner 2012). Evidence that the bilinguals' language systems are profoundly interactive has been reported for bilinguals whose two languages both share the same script and have different scripts (e.g. Gollan, Forster, and Frost 1997; Hoshino and Kroll 2008; Poarch and Van Hell 2014; Schwartz and Van Hell 2012; Thierry and Wu 2007), and even in sign-speech bilinguals who use languages from

https://doi.org/10.1515/9783110435351-018

different modalities (e.g. Van Beijsterveldt and Van Hell 2012; Morford et al. 2014). Current theoretical models depicting the bilingual mental lexicon depart from the assumption of language non-selectivity.

Models of bilingual lexical processing describe the orthographic, phonological, and semantic representations of words in the mental lexicon, and the consequences for lexical processing. These models can be classified into two categories. One class focuses on the architecture of the bilingual mental lexicon, in particular, the mappings between lexical and semantic information, and how these mappings change in the course of L2 learning (e.g. De Groot 1992; Van Hell and De Groot 1998; Kroll and Stewart 1994). The other describes the activation of lexical information, how orthographic, phonological, and semantic codes in the bilinguals' two languages interact during lexical processing, and how these cross-language interactions change with increased proficiency in the L2 (Dijkstra and Van Heuven 2002; Grainger, Midgley, and Holcomb 2010). Given our emphasis on cross-lexical processing and cross-language interactions, we focus on the second type.

The Bilingual Interactive Activation Plus model (BIA+; Dijkstra and Van Heuven 2002) is a localist-connectionist model postulating that the visual presentation of a word leads to parallel activation of orthographic codes in L1 and in L2, depending upon the similarity with the presented word and the codes' resting level activation. The activated orthographic representations then activate associated semantic and phonological representations, and from the interactions (resonance) among these three codes, the lexical code corresponding to the visual input emerges and is recognized. Recently, Shook and Marian (2013) described a variant of this model for the comprehension of speech, the Bilingual Language Interaction Network for Comprehension of Speech (BLINCS). Grainger, Midgley, and Holcomb (2010) describe a developmental variant, the developmental Bilingual Interactive Activation model (d-BIA). The d-BIA model incorporates into the Interactive Activation framework the idea that the strength of the connections between L2 orthographic-phonological codes and semantic codes are weak early in L2 learning, and become stronger as L2 proficiency increases (as in the Revised Hierarchical Model), and that translation pairs have overlapping and non-overlapping semantic features (as in the distributed feature model, e.g. Van Hell and De Groot [1998], or the sense model, e.g. Finkbeiner et al. [2004]).

Next, we review studies that examine lexical activation in bilingual memory during word comprehension and word production and that manipulate the phonological, orthographic, and semantic overlap of words in the bilinguals' languages. We specifically focus on behavioral and electrophysiological studies that compare the processing of cognates, homophones, or homographs with processing control words. Numerous other tasks and paradigms, such as manipulations of cross-language phonological or orthographic neighborhood density, cross-language priming, and word comprehension using the visual world paradigm have been used to study lexical activation in bilingual memory; for reviews, see De Groot (2010), Grosjean and Li (2012), Van Hell and Tanner (2012), Tokowicz (2015), and Schwieter (2015).

3. Bilingual lexical processing: studies manipulating lexical-semantic overlap

To examine the relative contribution of phonological, orthographic, and semantic codes to lexical activation in bilingual memory, and the nature of cross-language interactions, studies often manipulate the similarity of words across languages. For example, the processing of cognates (i.e. words with a similar or identical orthography, phonology, and semantics across languages, such as the English-Spanish translations 'piano'-*piano* or 'palace'-*palacio*) is compared to that of noncognate control words (i.e. words that only share meaning across languages, such as 'apple'-*manzana*), and the processing of homographs (e.g. words with identical spelling across languages, such as the Spanish word *pan* meaning 'bread', but spelled identically as the English 'pan') and homophones (words with similar or identical phonology across languages, such as the Dutch word *eekhoorn* that sounds like English 'acorn' but means 'squirrel') are compared to control words without orthographic or phonological overlap. The key assumption in using these words to examine the co-activation of languages is that if cognates, homographs, or homophones are processed differently from control words (e.g. faster and more accurately), the representations of cognates, homographs, or homophones in the two languages were co-activated at some point during lexical processing (at the level of orthography, phonology, or meaning), which altered the time-course or ease of processing.

Studies overwhelmingly report that cognates are processed faster and more accurately than noncognates across a wide variety of comprehension and production tasks, languages, and types of bilingual speakers (Van Hell and Tanner 2012). For example, adult multilinguals (Van Hell and Dijkstra 2002) and Dutch child L2 English learners (Brenders, Van Hell, and Dijkstra 2011) are faster to recognize cognates than noncognates in a lexical decision task measuring visual word recognition (Lemhöfer et al. 2008). Word production studies have found that pictures denoting cognates are named faster than pictures denoting noncognates in adult (Costa, Caramazza, and Sebastián-Gallés 2000) and child (Poarch and Van Hell 2012) bilinguals. Even though most studies used nouns as their critical materials, cognate effects have also been observed in verbs (Bultena, Dijkstra, and Van Hell 2013; Bultena, Dijkstra, and Van Hell 2014; Van Hell and De Groot 1998). Finally, studies have reported differences in the processing of homographs and homophones relative to control words (e.g. De Groot, Delmaar, and Lupker 2000; Dijkstra, Grainger, and Van Heuven 1999; Dimitropoulou, Duñabeitia, and Carreiras 2011; Durlik et al. 2016; Jared and Szucs 2002; Lagrou, Hartsuiker, and Duyck 2013; Macizo, Bajo, and Martín 2010; Studnitz and Green 2002), although the direction of the effect (facilitation or inhibition) depends on factors such as task demands and stimulus list composition (i.e. the proportion of homographs or homophones relative to controls, e.g. Dijkstra, Grainger, and Van Heuven 1999; Van Heuven et al. 2008). Moreover, several studies have found that the magnitude of cognate, homograph, and homophone effects is modulated by the degree of orthographical, phonological, and semantic overlap, both in same-script (e.g. Comesaña et al. 2015; Dijkstra, Grainger, and Van Heuven 1999; Lemhöfer and Dijkstra 2004; Van Hell and De Groot 1998) and different-script (e.g. Allen and Conklin 2013; Nakayama et al. 2014) bilinguals.

Effects of cross-language activation are not restricted to the processing of isolated words, but can also occur when words are embedded in a meaningful sentence context

(Lauro and Schwartz 2017). For example, Cop et al. (2017) recorded bilinguals' eye movements when they were reading a novel in their L1 or L2, and observed cognate facilitation effects on various eye movement measures in both L1 and L2. This suggests that even when reading connected and semantically rich discourse (i.e. a novel), lexical access in bilingual memory is not restricted to the target language.

Event Related Potentials (ERP) studies have also observed that variations in cross-language orthographic, phonological, or semantic overlap affect lexical processing, typically reflected in variations in the amplitude of the N400. The N400 is a large-amplitude negative-going ERP component prominent over centro-parietal scalp regions, beginning about 300 ms post-stimulus and reaching its maximum around 400 ms post-stimulus, and is considered to index lexical-semantic integration. Specifically, studies asking bilinguals to read cognates, homophones and control words have found a reduction in N400 amplitude in response to cognates in comparison to control words (Midgley, Holcomb, and Grainger 2011; Peeters, Dijkstra, and Grainger 2013), and to interlingual homophones in comparison to control words (Carrasco-Ortiz, Midgley, and Frenck-Mestre 2012). Moreover, ERP studies in which bilinguals name pictures denoting cognates or noncognates observed that ERP waveforms for cognates were more negative-going than those for noncognates (Christoffels, Firk, and Schiller 2007), a difference that emerged as early as 190 ms post-stimulus presentation (Strijkers, Costa, and Thierry 2010).

Taken together, the results of behavioral and electrophysiological studies that compare lexical processing of cognates, homographs, or homophones with control words support the notion that lexical activation in bilinguals is non-selective in nature and that lexical-semantic codes in the two languages interact during lexical processing. These assumptions are at the base of Bilingual Interactive Activation models, such as BIA+, BLINCS, and d-BIA.

The global population of bilingual speakers is highly diverse in terms of languages, proficiency in the second language, and cognitive abilities. This raises the question of whether the nature of cross-language interaction during lexical processing varies as a function of individual variation in language proficiency and executive function among adult bilingual speakers, and if so, how. Although most studies investigating cross-language interactions during lexical processing did not explicitly examine the role of individual differences, we now review the growing body of studies testing how the nature of cross-language interaction in bilingual lexical processing is affected by differences in L2 proficiency, or differences in relative proficiency in a bilingual's L1 and L2 (Section 4), and variations in executive function (Section 5). Please note that in the bilingual literature, various definitions of proficiency are used; see Surrain and Luk (2019) for a review of how studies published between 2005−2015 report different facets of bilingualism, including proficiency, language learning history, use of language at home, and language of schooling.

4. Cross-language activation and second language proficiency

In less proficient L2 speakers, the connections between phonological, orthographic, and semantic information in the L2 lexicon are weaker ('weaker links hypothesis', Gollan et al. [2008]), which will lead to slower L2 processing. But weaker connections between

L2 phonological, orthographic, and semantic codes as compared to the corresponding codes in the typically stronger L1 also have implications for patterns of cross-language activation in bilingual memory. We now discuss empirical studies that examined how differences in L2 proficiency between groups of participants, or differences in relative proficiency in a bilingual's L1 and L2 (using a within-participant design), affect parallel activation of the two languages during word recognition and word production. We focus on studies that manipulated lexical overlap across languages (i.e. by studying cognates, homophones, homographs; for other tasks, see Van Hell and Tanner 2012).

In one of the first word recognition studies that manipulated proficiency within participants, Van Hell and Dijkstra (2002) tested Dutch-English-French trilinguals, who all spoke Dutch as their L1, English as L2, and French as L3 (confirmed by proficiency tests after the experiment). In a lexical decision task, the trilinguals were presented with words in their L1 Dutch that were either cognates with L2 English, cognates with L3 French, or noncognates. The trilinguals' performance showed a cognate facilitation effect for both L1 Dutch-L2 English cognates, but not for the Dutch words that were cognates with French, their weakest language. In a second experiment, a new group of Dutch-English-French trilinguals were recruited whose proficiency in L3 French was higher than that of the first group (L1 Dutch and L2 English proficiencies were similar across the two groups), and were presented with the same materials. The trilinguals with higher L3-proficiency now demonstrated a cognate facilitation effect for both the L1 Dutch-L2 English cognates and the L1 Dutch-L3 French cognates. This shows that knowledge of a weaker language can influence processing in the native language. It also indicates that reading words in the native L1 co-activates the nontarget L2 and L3, but that a certain level of non-target language proficiency is needed for these co-activated systems to affect L1 lexical processing (and for a cognate facilitation effect to emerge).

Brenders, Van Hell, and Dijkstra (2011) confirmed that L2 proficiency affects the magnitude of cognate effects in lexical processing, but they also observed that the cognate facilitation effect is less robust in less proficient L2 speakers, and more vulnerable to contextual factors (here: stimulus list composition). In a series of lexical decision experiments, they presented cognates and noncognates in L1 and L2 to native Dutch children who were beginning (5th and 6th graders) or more advanced (7th and 9th graders) classroom learners of L2 English. All four groups of L2 learners recognized cognates faster than noncognates in L2, but not in L1, corroborating earlier findings that cognate effects only emerge when proficiency in the nontarget language is sufficiently high. They then tested the stability of the cognate facilitation effect in L2 lexical decision by adding homographs and homophones to the list of cognates and noncognates, and presented these words to new groups of beginning and more advanced L2 learners. Remarkably, all learners now recognized cognates *slower* than noncognate controls; homographs and homophones were also recognized slower than noncognate controls. The mixed presentation of cognates, homographs, and homophones thus resulted in a cognate inhibition effect rather than a facilitation effect. This reversal of the cognate effect contrasts with findings in highly proficient adult and child bilinguals (e.g. Dijkstra, Grainger, and Van Heuven 1999; Schröter and Schroeder 2016), who demonstrated robust cognate facilitation effects in conditions where cognates were mixed with homographs and homophones ('false friends'), or were not mixed. Brenders, Van Hell, and Dijkstra's (2011) findings thus show that the cognate effect in low and moderately proficient L2 learners is sensitive to the composition of the list the critical words are embedded in, indicating that the

nature of cross-language activation may be more sensitive to contextual factors in low-to-moderately proficient L2 learners than in proficient bilinguals, and that L2 learners and proficient bilinguals resolve lexical-semantic ambiguity in different ways. Specifically, Brenders, Van Hell, and Dijkstra (2011) argue that when cognates are mixed with false friends, L2 learners are confused by word form ambiguity for both cognates and false friends; they realize that orthographic/phonological overlap alone is not a sufficient basis for their response, which slows down performance for cognates in the mixed list relative to cognates in the non-mixed list. Proficient bilinguals, however, use their better developed and stronger lexical-semantic links, and the co-activated semantic codes of two cognate readings, to resolve ambiguity, which facilitates cognate processing irrespective of the presence of false friends.

Word production studies have also found that relative language proficiency affects the co-activation of languages in bilingual memory, as indexed by the cognate facilitation effect (Costa, Caramazza, and Sebastián-Gallés 2000; Poarch and Van Hell 2012; Rosselli et al. 2014). For example, Rosselli et al. (2014) asked balanced and non-balanced Spanish-English bilinguals to name pictures, denoting cognates and noncognates, in Spanish and English. The balanced bilinguals demonstrated cognate effects of similar magnitude in the two languages, but the unbalanced bilinguals showed a larger cognate effect when naming in the nondominant language than in the dominant language.

The finding that bilinguals' relative language proficiency in the two languages affects cross-language activation has also been reported for sign-speech bilinguals, thereby extending this effect to bimodal bilinguals. Morford et al. (2014) asked two groups of sign-speech bilinguals, deaf American Sign Language-dominant and hearing English-dominant bilinguals, to judge the semantic relatedness of word pairs presented in English. Critically, a subset of the semantically related and unrelated English word pairs had phonologically related translations in American Sign Language (ASL). In the ASL-dominant bilinguals, the English word pairs activated their ASL translation equivalents, speeding up performance when semantically related English word pairs had similar form translations in ASL (facilitation), but slowing down performance when semantically unrelated words had similar form translations in ASL (interference). In the hearing English-dominant bilinguals, English word pairs also activated their ASL translation equivalents, but the effects were smaller and less consistent.

In sum, adult and child bilinguals and multilinguals can show bi-directional cognate effects in L1 and L2 lexical processing, provided that proficiency in the nontarget language is sufficiently high. Moreover, cognate facilitation effects decrease as a function of proficiency, and are typically smaller in L1 than in L2 word recognition and word production, which indicates that the co-activation of target and nontarget language information is modulated by variation in bilinguals' language proficiency.

5. Cross-language activation and executive function

In recent years, cross-language interaction during bilingual lexical processing has been linked with cognitive control and executive function. Re-analyzing existing datasets, Linck, Hoshino, and Kroll (2008) found that bilinguals who had better inhibitory control showed smaller cognate effects in an L2 picture naming task, and bilinguals who had

greater working memory capacity showed smaller cognate effects in an L2 word naming task (no effects were found for working memory capacity in picture naming or inhibitory control in word naming). These re-analyzed datasets suggest that bilinguals with better inhibitory control and/or greater working memory capacity may be better able to limit lexical activation to the target language and curtail non-target language influences.

Recent studies explicitly investigated how individual differences in cognitive control and executive function modulate bilinguals' susceptibility to cross-language activation during lexical processing in same-script bilinguals (Blumenfeld and Marian 2013; Mercier, Pivneva, and Titone 2014; Pivneva, Mercier, and Titone 2014), different-script bilinguals (Chen et al. 2017; Prior et al. 2017), and bimodal bilinguals (hearing sign-speech bilinguals, Giezen et al. 2015). In a visual world eye-tracking study, Blumenfeld and Marian (2013) had English-Spanish bilinguals listen to English words (e.g. 'comb') and identify corresponding pictures from a display that also included pictures of Spanish phonological competitors (e.g. *conejo*, meaning 'rabbit'). Participants also completed a nonlinguistic spatial Stroop task measuring inhibitory control and conflict resolution. In the analyses, the authors related individual differences in inhibitory control to the degree and time-course of cross-language competition as measured in the visual world eye-tracking task (correlating performance on the executive function and lexical processing tasks is also used in studies to be discussed below). Bilinguals who showed better performance on the Stroop task also showed stronger co-activation of two languages during the early stages of word recognition (300–500 ms after word-onset), but reduced co-activation during later stages of word recognition. These findings indicate that the nature of the relationship between inhibitory control and co-activation of languages changes in the course of lexical comprehension: better inhibitory control abilities are associated with early and strong cross-language competitor activation, followed by efficient resolution of this cross-language competition in later stages of word recognition.

Using a similar visual world eye-tracking task, Mercier, Pivneva, and Titone (2014) also observed a relation between executive function and cross-language competitor activation in French-English bilinguals, but in the opposite direction to Blumenfeld and Marian (2013): increased inhibitory control abilities were associated with less cross-language competition. The direction of Mercier, Pivneva, and Titone's (2014) result pattern is paralleled by Pivneva, Mercier, and Titone (2014), who had French-English bilinguals read L2 English sentences that contained interlingual homographs, cognates, or control words while their eye movements were recorded, and complete a battery of executive function tasks. Bilinguals with greater executive control showed reduced co-activation of languages, as indexed by a smaller homograph interference effect (executive control was not related to the magnitude of the cognate facilitation effect).

Testing Arabic-Hebrew bilinguals whose languages do not share an orthographic system, Prior et al. (2017) examined whether bilinguals' executive control and L2 proficiency impact their ability to overcome interference from the L1 during L2 processing. By studying different-script bilinguals, they also address the question whether associations between executive function and cross-language activation observed in same-script bilinguals (Blumenfeld and Marian 2013; Mercier, Pivneva, and Titone 2014; Pivneva, Mercier, and Titone 2014) originate from shared orthographic systems, or reflect mechanisms for resolving cross-language interference that characterize all bilinguals, including different-script bilinguals. Lexical interference was measured using a cross-modal semantic similarity judgment task in L2 Hebrew. In critical trials, bilinguals listened to an

interlingual homophone in Hebrew (e.g. 'bread' in Hebrew is pronounced as *leh'em* and sounds like the Arabic word *lah'em* meaning 'meat'), followed by a visually presented Hebrew word related to the meaning of the homophone (e.g. 'butcher store') in the non-target language (here, Arabic). If presentation of the first L2 Hebrew word (interlingual homophone) co-activates the non-target language, the semantic similarity judgment of the word pair is predicted to slow down (lexical interference), which was indeed what Prior et al. (2017) observed. However, the lexical interference effect did not correlate with executive function abilities, or with L2 proficiency, which contrasts with earlier studies of same-script bilinguals by Blumenfeld and Marian (2013), Mercier, Pivneva, and Titone (2014), and Pivneva, Mercier, and Titone (2014), who did find a relation between the co-activation of languages during lexical processing and executive function (albeit in different directions). Moreover, suggestive electrophysiological evidence for such a relation in different-script Korean-English bilinguals is reported by Chen et al. (2017), who found that a smaller P300 effect of incongruent versus neutral trials in a nonlinguistic flanker task (indexing cognitive control) was positively associated with a smaller LPC effect of homophone related versus unrelated word pairs in a linguistic semantic relatedness task; Chen et al. (2017) take the P300 to index cognitive effort and attentional resources, and the LPC to index the consequence of exerting language inhibition.

Giezen et al. (2015) examined whether bimodal bilinguals engage inhibitory control when resolving cross-language competition from the nontarget language during lexical processing of languages in different modalities, ASL and English. Using a visual world eye tracking paradigm, highly proficient hearing ASL-English bilinguals heard a word (e.g. 'chair') and identified corresponding pictures from a display that also included pictures of cross-linguistic competitors (here: 'train', which is highly similar to the target word 'chair' in three out of four major sign parameters). Participants also completed a nonlinguistic spatial Stroop task measuring inhibitory control. If inhibitory control is recruited to resolve cross-language competition that originates at lexico-semantic levels of processing (beyond merely perceptual competition resulting from orthographic-phonological similarity in unimodal same-script bilinguals), sign-speech bilinguals should show an association between language co-activation and performance on the inhibitory control task. Giezen et al. (2015) indeed found that bimodal bilinguals with smaller Stroop effects (so better inhibitory control) experienced reduced competition from the nontarget language (ASL) during the early stages of auditory word recognition (180–260 ms post-target word onset). This pattern parallels findings in unimodal same-script bilinguals (Blumenfeld and Marian 2013; Mercier, Pivneva, and Titone 2014; Pivneva, Mercier, and Titone 2014).

To conclude, research examining whether individual differences in cognitive control modulate bilinguals' susceptibility to cross-language activation during lexical processing indicates that both unimodal (same-script) and bimodal bilinguals engage domain-general cognitive control processes to resolve cross-language competition between languages during word recognition. That these effects were obtained in both unimodal (same-script) and bimodal bilinguals suggests that inhibitory control mechanisms recruited to resolve cross-language competition are not restricted to phonologically-orthographically similar language systems, although further research is needed to elucidate this issue in different-script bilinguals (Chen et al. 2017; Prior et al. 2017). As Prior et al. (2017) point out, different measures of executive control do not always correlate (see also Poarch and Van

Hell 2019), so discrepancies between the studies in this literature may at least be partly related to the use of different executive control measures. Finally, whether better cognitive control abilities are associated with stronger (Mercier, Pivneva, and Titone 2014; Pivneva, Mercier, and Titone 2014) or weaker (Blumenfeld and Marian 2013; Giezen et al. 2015) co-activation of the nontarget language, and how the nature of this relation may change in the time course of lexical processing, remain questions for future studies.

6. Concluding remarks

Research on lexical processing in bilinguals has found ubiquitous evidence that the bilinguals' lexical systems are profoundly interactive, even when the social and linguistic context calls for only one language. We reviewed selected studies examining how individual variation in language proficiency and executive function impact cross-language activation during word comprehension and production. Current evidence indicates that the co-activation of target and non-target languages is modulated by variations in bilinguals' language proficiency, such that lexical processing in the target language is only affected by non-target language information if proficiency in the non-target language is sufficiently high. Furthermore, both unimodal (same-script) and bimodal bilinguals engage domain-general cognitive control processes to manage cross-language activation and resolve cross-language competition between languages during lexical processing. One of the next questions is whether language proficiency modulates how bilinguals engage inhibitory control when resolving cross-language competition between languages.

Acknowledgments

The writing of this paper was supported by NSF grants BCS-1349110, SMA-1514276, DUE-1561660, DUE IUSE-172681, and OISE-1545900 to Janet van Hell.

7. References

Allen, David & Kathy Conklin
2013 Cross-linguistic similarity and task demands in Japanese-English bilingual processing. *PLoS ONE* 8(8). 1–14.

Blumenfeld, Henrike & Vioria Marian
2013 Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology* 25(5). 547–567.

Brenders, Pascal, Janet G. van Hell & Ton Dijkstra
2011 Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of Experimental Child Psychology* 109(4). 383–396.

Bultena, Sybrine, Ton Dijkstra & Janet G. van Hell
2013 Cognate and word class ambiguity effects in noun and verb processing. *Language and Cognitive Processes* 28(9). 1350–1377.

Bultena, Sybrine, Ton Dijkstra & Janet G. van Hell
2014 Cognate effects in sentence context depend on word class, L2 proficiency, and task. *Quarterly Journal of Experimental Psychology* 67(6). 1214–1241.
Carrasco-Ortiz, Haydee, Katherine J. Midgley & Cheryl Frenck-Mestre
2012 Are phonological representation in bilinguals language-specific? An ERP study on interlingual homophones. *Psychophysiology* 49(4). 531–543.
Chen, Peiyao, Susan C. Bobb, Noriko Hoshino & Viorica Marian
2017 Neural signatures of language co-activation and control in bilingual spoken word comprehension. *Brain Research* 1665. 50–64.
Christoffels, Ingrid K., Christine Firk & Niels O. Schiller
2007 Bilingual language control: An event-related brain potential study. *Brain Research* 1147. 192–208.
Comesaña, Montserrat, Pilar Ferré, Jaoquin Romero, Marc Guasch, Ana P. Soares & Teófilo Garcia-Chico
2015 Facilitative effect of cognate words vanishes when reducing orthographic overlap: The role of stimuli list composition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 41(3). 614–635.
Cop, Uschi, Nicolas Dirix, Eva Van Assche, Dennis Drieghe & Wouter Duyck
2017 Reading a book in one or two languages? An eye movement study of cognate facilitation in L1 and L2 reading. *Bilingualism: Language and Cognition* 20(4). 747–769.
Costa, Albert, Alfonso Caramazza & Nuria Sebastián-Gallés
2000 The cognate facilitation effect: Implications for models of lexical access. *JEP: LMC* 26(5). 1283–1296.
De Groot, Annette M. B.
1992 Bilingual lexical representation: A closer look at conceptual representations. In Ram Frost & Leonard Katz (eds.), *Orthography, phonology, morphology, and meaning*, 389–412. Amsterdam: Elsevier.
De Groot, Annette M. B.
2010 *Language and cognition in bilinguals and multilinguals*. Hove: Psychology Press.
De Groot, Annette M. B., Philip Delmaar & Stephen J. Lupker
2000 The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 53A(2). 397–428.
Dijkstra, Ton & Walter J. B. van Heuven
2002 The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* 5(3). 175–197.
Dijkstra, Ton, Jonathan Grainger & Walter J. B. van Heuven
1999 Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language* 41(4). 496–518.
Dimitropoulou, Maria, Jon Andoni Duñabeitia & Manuel Carreiras
2011 Phonology by itself: Masked phonological priming effects with and without orthographic overlap. *Journal of Cognitive Psychology* 23(2). 185–203.
Durlik, Joanna, Jakub Szewczyk, Marek Muszynski & Zofia Wodniecka
2016 Interference and inhibition in bilingual language comprehension: Evidence from Polish-English interlingual homographs. *PLoS One* 11(3). e0151430
Finkbeiner, Matthew, Kenneth I. Forster, Janet Nicol & Kumiko Nakamura
2004 The role of polysemy in masked semantic and translation priming. *Journal of Memory & Language* 51. 1–22.
Giezen, Marcel R., Henrike K. Blumenfeld, Anthony Shook, Viorica Marian & Karen Emmorey
2015 Parallel language activation and inhibitory control in bimodal bilinguals. *Cognition* 141. 9–25.

Gollan, Tamar H., Kenneth I. Forster & Ram Frost
1997 Translation priming with different scripts: Masked priming with cognates and non-cognates in Hebrew-English bilinguals. *JEP: LMC* 23(5). 1122–1139.
Gollan, Tamar H., Rosa I. Montoya, Cynthia Cera & Tiffany C. Sandoval
2008 More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language* 58(3). 787–814.
Grainger, Jonathan, Katherine Midgley & Philip J. Holcomb
2010 Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In Michèle Kail & Maya Hickmann (eds.), *Language acquisition across linguistic and cognitive systems*, 267–284. New York: John Benjamins.
Grosjean, Francois & Ping Li
2012 *The psycholinguistics of bilingualism*. Wiley-Blackwell.
Hoshino, Noriko & Judith F. Kroll
2008 Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition* 106. 501–511.
Jared, Debra & Carrie Szucs
2002 Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition* 5(3). 225–239.
Kroll, Judith F. & Erika Stewart
1994 Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33(2). 149–174.
Lagrou, Eveline, Robert J. Hartsuiker & Wouter Duyck
2013 The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingualism: Language and Cognition* 16(3). 508–517.
Lauro, Justin & Ana I. Schwartz
2017 Bilingual non-selective access in sentence contexts: A meta-analytic review. *Journal of Memory and Language* 92. 217–233.
Lemhöfer, Kristin & Ton Dijkstra
2004 Recognizing cognates and interlingual homographs: Effects of code-similarity in language-specific and generalized lexical decision. *Memory and Cognition* 32(4). 533–550.
Lemhöfer, Kristin, Ton Dijkstra, Herbert Schriefers, R. Harold Baayen, Jonathan Grainger & Pienie Zwitserlood
2008 Native language influences on word recognition in a second language: A megastudy. *JEP: LMC* 34(1). 12–31.
Linck, Jared A., Noriko Hoshino & Judith F. Kroll
2008 Cross-language lexical processes and inhibitory control. *Mental Lexicon* 3(3). 349–374.
Macizo, Pedro, Teresa M. Bajo & Maria Cruz Martín
2010 Inhibitory processes in bilingual language comprehension: Evidence from Spanish-English interlexical homographs. *Journal of Memory and Language* 63(2). 232–244.
Mercier, Julie, Irina Pivneva & Debra Titone
2014 Individual differences in inhibitory control relate to bilingual spoken word processing. *Bilingualism: Language and Cognition* 17(1). 89–117.
Midgley, Katherine J., Philip J. Holcomb & Jonathan Grainger
2011 Effects of cognate status on word comprehension in second language learners: An ERP investigation. *Journal of Cognitive Neuroscience* 23(7). 1634–1647.
Morford, Jill P., Judith F. Kroll, Pilar Piñar & Erin Wilkinson
2014 Bilingual word recognition in deaf and hearing signers: Effects of proficiency and language dominance on cross-language activation. *Second Language Research* 30(2). 251–271.

Nakayama, Mariko, Rinus G. Verdonschot, Christopher R. Sears & Stephen J. Lupker
2014 The masked cognate translation priming effect for different-script bilinguals is modulated by the phonological similarity of cognate words: Further support for the phonological account. *Journal of Cognitive Psychology* 26(7). 714−724.
Peeters, David, Ton Dijkstra & Jonathan Grainger
2013 The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language* 68(4). 315−332.
Pivneva, Irina, Julie Mercier & Debrah Titone
2014 Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40(3). 787−796.
Poarch, Gregory J. & Janet G. van Hell
2012 Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology* 111(3). 419−438.
Poarch, Gregory J. & Janet G. van Hell
2014 Cross-language activation in same-script and different-script trilinguals. *International Journal of Bilingualism* 18(6). 693−716.
Poarch, Gregory J. & Janet G. van Hell
2019 Does performance on executive function tasks correlate? Evidence from child trilinguals, bilinguals, and second language learners. In Irina Sekerina, Virgina Valian & Lauren Spradlin (eds.), *Bilingualism, executive function, and beyond: Questions and insights* [Studies in bilingualism 57], 223–236. Amsterdam: John Benjamins.
Prior, Anat, Tamar Degani, Sehrab Awawdy, Rana Yassin & Nachshon Korem
2017 Is susceptibility to cross-language interference domain specific? *Cognition* 165. 10−25.
Rosselli, Mónica, Alfredo Ardila, Mária B. Jurado & Judy L. Salvatierra
2014 Cognate facilitation effect in balanced and non-balanced Spanish-English bilinguals using the Boston Naming Test. *International Journal of Bilingualism* 18(6). 649−662.
Schröter, Pauline & Sascha Schroeder
2016 Orthographic processing in balanced bilingual children: Cross-language evidence from cognates and false friends. *Journal of Experimental Child Psychology* 141. 239−246.
Schwartz, Ana I. & Janet G. van Hell
2012 Bilingual visual word recognition in sentence context. In James Adelman (ed.), *Visual Word Recognition*, 131−150. Hove: Psychology Press.
Schwieter, John W. (ed.)
2015 *The Cambridge handbook of bilingual processing.* Cambridge: CUP.
Shook, Anthony & Viorica Marian
2013 The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition* 16(2). 304−324.
Strijkers, Kristof, Albert Costa & Guillaume Thierry
2010 Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex* 20(4). 912−928.
Surrain, Sarah, & Gigi Luk
2019 Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005−2015. *Bilingualism: Language and Cognition* 22(2). 401–415.
Thierry, Guillaume & Yan Jing Wu
2007 Brain potentials reveal unconscious translation during foreign-language comprehension. *PNAS* 104(30). 12530−12535.
Tokowicz, Natasha
2015 *Lexical Processing and Second Language Acquisition.* New York: Routledge.

Van Beijsterveldt, Liesbeth M. & Janet G. van Hell
2012 Temporal reference marking in narrative and expository text written by deaf children and adults: A bimodal bilingual perspective. *Bilingualism: Language and Cognition* 15(1). 128–144.
Van Hell, Janet G. & Annette M. B. de Groot
1998 Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition* 1(3). 193–211.
Van Hell, Janet G. & Ton Dijkstra
2002 Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review* 9(4). 780–789.
Van Hell, Janet G. & Darren Tanner
2012 Second language proficiency and cross-language lexical activation. *Language Learning* 62(s2). 148–171.
Van Heuven, Walter J. B., Herbert Schriefers, Ton Dijkstra & Peter Hagoort
2008 Language conflict in the bilingual brain. *Cerebral Cortex* 18(11). 2706–2716.
Von Studnitz, Roswitha E. & David Green
2002 Interlingual homograph interference in German-English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition* 5(1). 1–23.

Janet G. van Hell, University Park, PA (USA)
Katharine Donnelly Adams, University Park, PA (USA)
Fatemeh Abdollahi, University Park, PA (USA)

19. Metalinguistic awareness and multilingual development

1. Introduction
2. Defining ‘meta-linguistic’
3. Researching metalinguistic awareness in two and more languages in contact
4. Exploring multilingual awareness as an emergent quality of the multilingual system
5. Focus on metacognition in future research
6. References

1. Introduction

The definition of the role that metalinguistic competences play in multilingual learning and use presents a challenge to research on language contact and language learning in general, and to the theoretical and methodological paradigms that guide such research in particular. This chapter first defines metalinguistic competences by relating them to abilities and skills (Section 2), and then surveys how metalinguistic awareness is dealt with in L1, L2, and L3 learning research (Section 3). The chapter then presents Herdina and Jessner’s (2002) Dynamic Model of Multilingualism (DMM), and focuses on the crucial role that emergent properties of the multilingual system, such as metalinguistic competences, have been assigned in the DMM (Section 4). The brief outlook discussion

https://doi.org/10.1515/9783110435351-019

(Section 5) asks for a stronger focus on metacognition in research on language contact in multilingual development and education.

2. Defining 'meta-linguistic'

'Metalinguistic awareness' is a notion that is often used interchangeably with other notions, such as 'knowledge about language' and 'language awareness', that are, in fact, broader in scope. The Association for Language Awareness, for example, defines language awareness on its website as "explicit knowledge about language, and conscious perception and sensitivity in language learning, language teaching and language use" (see www.languageawareness.org). Cots and Garrett (2018: 3) observe that other authors on the website of the same association also include unconscious/nonconscious phenomena in the discussion of language awareness. They also state that even though the core business of language awareness still appears to be concerned with (first, second, and third) language learning and teaching (van Essen 2008: 3), language awareness has been studied in "areas ranging from professional communication, translation, visual communication to other areas such as social psychology of language, sociolinguistics, dialectology and the sociology of language" (Cots and Garrett 2018: 5). To the multidisciplinary approach to language awareness, one can add the existence of different research traditions in different countries, especially in the educationally-oriented field of language awareness. Next to the language awareness movement in the UK that is based on the work of Hawkins (1984), France and French-speaking Canada have witnessed the emergence of a movement called *éveil aux languages* promoted in the work of Candelier (see Troncy 2014). In language education in the German-speaking world, for instance, the label *Begegnung mit Sprachen* was used (Gnutzmann 1997). Differences in research foci are also linked to different meanings of 'awareness' and related notions in different languages. The two interchangeable terms in Italian *consapevolezza* ('awareness') and *coscienza* ('consciousness') are not treated as synonyms in English. In French *conscience* and *prise de conscience* describe notions of a process, and the German *Sprachbewusstsein* as the most common term used for 'language awareness' could also be substituted by *Sprachbewusstheit* (Jessner 2006: 41).

When it comes to the metalinguistic dimension of language awareness, a dimension that narrows its scope, Simard and Guitiérez (2018: 208–209) distinguish between metalinguistic knowledge, awareness, reflection, activity, and ability. Their synthesis reads as follows:

> Prevalent features in each definition of the metalinguistic constructs […] can be identified. First, metalinguistic knowledge often refers to the explicit, conscious representations of language, whereas metalinguistic ability has to do with the capacity to consciously manipulate linguistic elements. While both metalinguistic awareness and metalinguistic reflection (and activity) seem to deal with attention to language, the former has to do with the individual's ability to attend to linguistic elements and the latter to cognitive processes involved in actually doing so. Interestingly, some authors' definitions of metalinguistic awareness include the term "ability" (e.g. Dillon 2009; Jessner 2008; Reder et al. 2013). […] [M]etalinguistic ability has been commonly defined as the individual's ability to manipulate linguistic elements, a definition which has also been used to refer to metalinguistic awareness.

In this chapter, the focus is on metalinguistic awareness. Most scholars seem to agree that metalinguistic awareness can be described as the ability to focus on linguistic form and to switch focus between form and meaning (Simard and Guitiérez 2018). In other words: individuals who are metalinguistically aware are able to categorize words into parts of speech, switch focus between form, function, and meaning, and explain why a word has a particular function. Consequently, the distinction between explicit learning and implicit learning is linked to the development of levels of metalinguistic awareness (see also below).

3. Researching metalinguistic awareness in two and more languages in contact

Some of the scholars whose work focuses solely on monolingual children have contributed considerably to our present understanding of metalinguistic awareness. Gombert (1992: 13), for example, views metalinguistic activities as a subfield of metacognition. He defines metalinguistic activities as activities of reflection on language and its use, as well as the subjects' ability to intentionally monitor and plan their own methods of linguistic processing both in comprehension and production. That emergent metalinguistic abilities form the reflection of underlying changes in cognitive abilities was already pointed out by Piaget. His point of view was taken up by Karmiloff-Smith (1992), who describes the emergence of conscious access to the mind via metalinguistic abilities as a cyclical process by which information already present in the organism is made progressively available via redescriptive processes to other parts of the cognitive system. Already Vygotsky ([1934] 1986: 160) stated that "… a child's understanding of his native language is enhanced by learning a foreign".

Pinto, Titone, and Trusso (1999) paid particular attention to the development of metalinguistic abilities in children, adolescents, and adults. The Italian researchers around Renzo Titone developed the most comprehensive test of metalinguistic abilities, which actually consists of three sets of metalinguistic awareness tests (MATs): MAT1 for children between four and six, MAT2 for children between nine and fourteen, and MAT3 for adolescents and adults (the tests have been translated into various European languages, such as English, French, Spanish, and German; see the special issue of *Rivista di psicolinguistica applicata* edited by Pinto [2015]). When the test battery − which originally had been developed to test metalinguistic abilities in Italian monolingual children − was used in the bilingual context of South Tyrol, it turned out that the bilingual subjects significantly outperformed their peers. This again makes clear that work on monolinguals has significantly contributed to current knowledge of metalinguistic awareness in more than one language. In the next sections, however, we turn to work on L2 and L3 acquisition.

3.1. Second language acquisition (SLA) and bilingualism

As noted by Simard and Guitérrez (2018: 209, 215), studies on metalinguistic constructs in the field of SLA over the last 30 years can be categorized into three types: (1) *type a*,

which concentrates on the relationship between metalinguistic constructs and learning success, (2) *type b*, which deals with the description of the nature and use of metalinguistic constructs, and (3) *type c*, which is concerned with the development of metalinguistic constructs among L2 learners. The measurement instruments used in SLA seem to rely on mainly two types of data: verbalization data and performance data.

In a number of SLA studies, grammaticality judgment tests have been widely used to elicit metalinguistic data as predictors of success or failure in language learning (Chaudron 1983). As explained by Birdsong (1989: 29), the concept of metalinguistic performance in L1 and L2 acquisition "applies both to the output from the learner and input to the learner. When metalinguistic performance is considered from the output side, the emphasis is on producing synonymity and grammaticality judgments, pointing out ambiguity, locating errors, explaining word choice, etc. From the input side, metalinguistic performance involves attending to formal features of the linguistic environment […]".

In more recent work, the nature of metalinguistic constructs has been at the center of interest. In Leow, Johnson, and Zarate-Sandez (2011), the focus is directed towards levels of awareness, i.e. noticing, reporting, and understanding. The authors argue that differentiation and selectivity in multilingual production seem to be governed by different levels of awareness, which clearly calls into question a bipolar discussion of the various phenomena. According to Leow (2015), the three main questions that have been discussed concern the role of awareness in language learning, the measurement of awareness during learning, and the interface between implicit and explicit learning (see Ellis [2011] for a review).

Bialystok's famous work on bilingual children that dates back to the 1990s has more recently been supplemented with her studies on bilingual processing in adults. Bialystok et al. (2004) suggest that lifelong bilingualism protects older adults from cognitive decline, as in Alzheimer's decease. A number of scholars have applied Bialystok's model of analysis and control as the metalinguistic dimensions of bilingual proficiency to investigate the impact of bilingualism on cognitive skills (e.g. Modirkhamene 2008). Analysis of linguistic knowledge is described as the skill component responsible for making explicit those representations that had previously been implicit or intuitive. Control of linguistic processing is described as the ability to selectively attend to specific aspects of a representation, particularly in misleading situations. Bialystok (e.g. 2001) concludes from her studies that there are no universal advantages for bilinguals, but that high levels of proficiency in both languages lead to advantages on tasks requiring more analyzed linguistic knowledge, as also indicated by Mohanty (1994) in his study of Kond tribal children in India. In a more recent study focusing on immersion programs (Bialystok and Barac 2012), it was revealed that level of proficiency in the testing language was linked to performance on metalinguistic tasks. Performance on executive control tasks was related to length of time in the immersion program, i.e. a distinction was found between representational structure and executive control. From a lifespan perspective "the bilingual profile for executive control in both children and adults emerges with experience in a bilingual environment" (Bialystok and Barac 2012: 71). Although recently the so-called 'bilingual advantage' has been questioned (de Bot 2017), evidence stemming from studies carried out in the context of third language learning (see Section 3.2) nevertheless seem to support the multilingual advantage perspective.

3.2. Third language acquisition (TLA) and multilingualism

For a considerable amount of time, the phenomenon of metalinguistic awareness was mainly studied by scholars with either a second language acquisition (SLA) or a bilingualism background. Studies on metalinguistic awareness in SLA for a long time mainly concentrated on the outcomes of grammaticality judgement tests. Work on bilingualism has traditionally studied a much wider range of metalinguistic skills, and has also linked metalinguistic awareness to cognitive advantages, such as communicative sensitivity, flexibility, and creativity, that are stimulated as a result of an individual's contact with two or more languages. Therefore, advantages develop not only in terms of multiple language competence, but also in terms of cognitive and social behavior, as discussed in Cook's (1991) construct of multicompetence, which explicitly refers to metalinguistic awareness as developed on a higher level in second language users. Already Hamers and Blanc (1989) argued that cognitive advantages seem to occur mainly in terms of higher creativity and reorganization of information. These findings on bilingual advantages have marked the history of research on bi- and multilingualism considerably. After periods of both negative and enthusiastic attitudes towards bi- and multilingualism (for a detailed discussion, see Jessner 2006), we have now arrived at a more sophisticated understanding of the complexity of the phenomenon. In the study of TLA, that understanding is enhanced through the fact that TLA research is heavily intertwined with research on multilingualism in general. Examples are provided in the following paragraphs.

In a number of studies of a young trilingual (Spanish/Catalan/English) child, Safont-Jordà (2013) reports the emergence of requestive behavior in the three languages. She concludes that the child's multilingual development of pragmatic skills shows features that are not found in bilingual children. This supports other studies (Montanari 2009) that also detected pragmatic flexibility and differentiation apparent in very young trilingual learners.

Creativity also seems to be increased in bi- and multilingual speakers. Kharkurin (2012) showed that certain cognitive mechanisms are encouraged by an individual's cross-linguistic and cross-cultural experiences, one of them being creativity. Other research also claims that there is a link between individuals' multilingualism and creativity, in the sense that multilingualism broadens access to information, and offers alternative ways of organizing thoughts and of perceiving the surrounding world (Furlong 2009).

Several areas of research have dealt with learning and development of three or more languages, which at the same time have focused on metalinguistic awareness in learning. These include studies on the effect of bilingualism on further language learning and on artificial language learning. Inherent in both research areas is an interest in language learning strategies. Nation and McLaughlin (1986) and Nayak et al. (1990) studied the learning of an artificial miniature linguistic system. Both studies pointed towards a learning advantage for expert foreign language learners over less experienced learners. Later on, Kemp (2001) found that the performance of multilingual learners on tests of grammaticality depended on the number of languages they knew. As a group, multilinguals turned out to perform better on explicit than implicit portions of metalinguistic tests. All these authors suggested that the multilinguals were more capable of adapting their strategies to the requirements of the task. This leads to the conclusion that one reason for the superior performance of multilingual participants is their greater flexibility in switching

strategies. More recently, this result was supported by two large-scale Greek studies (Mitits 2015; Psaltou-Joycey and Kantaridou 2009).

4. Exploring multilingual awareness as an emergent quality of the multilingual system

As briefly mentioned in the introduction, dynamic systems or complexity theory (DSCT) presents a new approach to the phenomenon of multilingualism. The study of TLA and trilingualism plays an important part in the exploration of multilingualism, as it is considered absolutely necessary to move beyond the study of two languages in contact (Flynn, Foley, and Vinnitskaya 2004).

As explained in DSCT, key features of a complex system are sensitivity to initial conditions and development of emergent properties through the interaction between systems. Sensitivity to initial conditions is a key feature of complex systems inherent to multilingualism. As pointed out by Aronin and Jessner (2015), this feature materializes in the problems related to language learning in a school context, in linguistic minority issues, and in new languages or new linguistic varieties in a society. The discussion of sensitivity to initial conditions is also known as the 'butterfly effect', which was modelled by Edward Lorenz for the study of meteorology and demonstrates that noticeable changes can be brought about by the very slightest change. We can find a good illustration of sensitivity to specific factors in studies dealing with L2 status in TLA.

In a number of studies, it turned out that in L3 production, the speakers did not, as expected, automatically rely on their L1, but also on their L2. In several studies of learning an L3 of Indo-European origin, it could be shown that L3 learners whose L1 is typologically unrelated to the L2 and/or the L3 tend to transfer knowledge from their L2, or in the case of bilinguals, from the related L1. The activation of prior language knowledge is influenced by factors such as psychotypology (perceived linguistic distance between languages), recency of use, level of proficiency in the target language, the foreign language effect, the tendency in language learners to activate an earlier second language in L3 performance, and the learner's perception of correctness of a target word (for an overview, see De Angelis, Jessner, and Kresic 2015).

In the Dynamic Model of Multilingualism (DMM), crosslinguistic interaction in multilinguals not only covers transfer and interference, and codeswitching and borrowing, but also another set of phenomena. These are non-predictable cognitive effects of the development of the multilingual which determine the development of the systems themselves. This development is referred to as 'paradox of transfer'. Such a view is also related, yet not identical, to Common Underlying Proficiency in the work of Cummins (1991) and Common Underlying Conceptual Base in Kecskes and Papp (2000).

Multilingual proficiency is defined as the dynamic interaction between the various psycholinguistic systems (LS_1, LS_2, LS_3, LS_n) in which the individual languages (L1, L2, L3, Ln) are embedded, crosslinguistic interaction, and what is called the M(ultilingualism)-factor. The latter refers to all the effects in multilingual systems which distinguish a multilingual from a monolingual system. It considers all those qualities that develop in a multilingual speaker/learner due to the increase in language contact(s) and depend on the perceived communicative needs of the individual. In other words: the

constantly changing psycholinguistic systems of the multilingual individual interact with each other in a non-additive, but cumulative, way.

In DMM, metalinguistic awareness has been identified as one, or in fact, *the* key factor in multilingual learning (Jessner 2006, 2008). It considers multilingual acquisition processes to be supported and enhanced by robust synergetic effects which emerge in experienced multilingual learners on account of the extensive (meta)linguistic and metacognitive knowledge which they accumulate as language learners and users. The awareness of differences and similarities between linguistic knowledge in the languages in contact seems to increase with the number of languages involved. In addition, in experienced language learners, acquisition processes are facilitated by the availability of an enhanced (multilingual) monitor that supervises and surveys all language activity and processing, and by elevated levels of metalinguistic awareness and ability. Metalinguistic awareness and abilities are considered characteristic features of multilingual speakers' multi(ple) competence, and central components of the so-called M-factor, which accounts for their superiority in various domains of verbal and non-verbal performance.

As already mentioned, the key factor in the M-effect − as it might also be referred to, since it is difficult to decide whether it constitutes a precondition for or a result of multilingualism − is metalinguistic or multilingual awareness. The domains of the multilingual system in which metalinguistic ability and awareness play a crucial role are linguistic development in general, the development of cognitive, metacognitive, and information-processing abilities, and the development of literacy skills.

Emergent properties are the outcome or products of interactions between the parts of a system, but they are not only the sum of their parts. Emergent phenomena acquire properties that are different from the properties of their parts, and they all show unpredictable behavior. According to the DMM, emergent phenomena are specific skills in a trilingual which develop beyond knowledge of the three languages. Among those are certain social skills, such as communicative sensitivity and pragmatic flexibility, which also seem to develop to a higher degree in a bi- and/or multilingual system (see above).

In a number of studies on both the primary and secondary educational level carried out by the DyME group at Innsbruck University (DyME stands for Dynamics of multilingualism with English; www.dyme.uibk.ac.at), research has focused on the exploration of multilingual awareness within strategic processing and techniques, as well as on the nature of crosslinguistic interaction in multilingual development. It turned out that multilingual learners are specific language learners with a significant advantage in the development of metalinguistic skills in comparison with bilinguals or L2 users. This has been evidenced in two, obviously differing sociolinguistic contexts, Tyrol and South Tyrol, where learners developed an enhanced level of multilingual awareness based on various multilingual teaching approaches. Whereas in North Tyrol German is the official language, South Tyrol (the most Northern part of Italy) has three official languages: German, Italian, and Ladin.

The work of the DyME research group is based on both the DMM as the theoretical framework, and Jessner's work on linguistic awareness in trilingual students from South Tyrol studying English at Innsbruck University. In her introspective study on lexical search in third language production, Jessner (2006) found that crosslinguistic awareness and metalinguistic awareness, tested in the form of explicit metalanguage, exerted influence on the activation of individual languages in the multilingual mental lexicon (see also Jessner, Megens, and Graus 2016). In her work, she has been concerned with the

relationship between crosslinguistic interaction (that is, the activation of languages other than the target language during third language production) and metalinguistic awareness as key variables forming part of the M(ultilingualism)-factor. Crosslinguistic awareness in multilingual production is described as (a) tacit awareness shown by the use of cognates in the supporter languages (mainly in the use of combined strategies), and (b) explicit awareness in the case of switches that are introduced by meta-language. James (1996: 139) defined crosslinguistic awareness as knowledge held at the explicit (declarative) level of metacognition. In his crosslinguistic approach to language awareness he pointed out that the "*language transfer* issue of classical Contrastive Analysis becomes a new issue of *metalinguistic transfer* − its relationship to cross-linguistic awareness" (James 1996: 143, italics in original).

In two large-scale studies on linguistic awareness in language attrition carried out in both Tyrolean and South Tyrolean contexts (LAILA and LAILA-BICS), young adult students were asked to produce introspective think-aloud protocols during the decoding of a Romanian text, a language hitherto unknown to the students. At the time of the study, the subjects in both contexts had been in contact with at least 3 languages (German, English, Latin, and/or Italian or French) during their school career.

The examples from the LAILA and LAILA-BICS studies give extensive proof of the emergent properties within strategic processing when dealing with an unknown foreign language in a complex multilingual system. The examples evidence both the study participants' metalinguistic and their crosslinguistic awareness based on language typology and grammatical awareness, as well as language transfer informed by the use of supporter languages such as German, English, French, Italian, and Spanish, as well as Latin (Jessner, Török, and Pellegrini 2018). The strategies applied evidently differ from mono- and bilinguals, but to a certain extent also from less-experienced multilingual language learners, not only in terms of quantity, but also very much in their quality, therefore demonstrating the realm of the M-factor (Jessner and Török 2017).

In a different line of research, the DyME-group has focused on the application of the DMM in education in both Austria and South Tyrol. Traditional foreign language teaching and learning in (European) educational contexts is connected with the idea that students should "get as close as possible to monolingual native speakers" (Cook 2002: 239), a goal implicitly stated in the Common European Framework of Reference (Council of Europe 2001). The benchmark is that of native speaker competence in all domains coupled with an implicit assumption of linear development across the life span, a conceptual tenet that stands in clear opposition to the ideas put forward in the DMM and other related or supportive approaches to multilingual (or rather: plurilingual) teaching that have also been developed in France and Canada (Moore 2006; Young and Helot 2003; Cummins [2017] on teaching for transfer).

Jessner, Allgäuer-Hackl, and Hofer (2016) present a multilingual and multicompetence approach to teaching based on DMM. The examples and classroom activities described have been observed or used in multilingual classes in South Tyrol and Austria with students of different age groups, varying combinations of languages spoken and learned, and varying levels of linguistic competencies. Multilingual activities can be designed for multilingual classes, but also adapted to foreign language classes in such a way that they support the development and use of the target language. They can also, as some examples will show, be integrated into subject teaching in general and will, as we claim, contribute to more efficient (language) learning and teaching. Multilingual ap-

proaches motivate students to develop more language learning strategies and expand the types of strategies used (see also Dahm 2015).

This also ties in with the notion of a qualitative change in language learning skills in trained multilinguals, as discussed by Hofer and Jessner (2016). In a study of primary school children in South Tyrol, they note that regular and systematic reference to structural phenomena and language forms in all three languages serves to reinforce the pupils' meta- and cross-linguistic awareness, and to enable them to perceive and exploit potential synergies.

5. Focus on metacognition in future research

The primary goal of this entry is to look into the nature of metalinguistic awareness and identify it as an emergent property in multilingual learners and users. At the same time, it describes metalinguistic awareness as a feature of multilingual proficiency that cannot be studied with a monolingual yardstick. The role that metalinguistic competences and awareness of such competences plays in research on language learning has been the focus lately in more work on metacognition in language learning (and learning in general). This ability to reflect on one's own thoughts, knowledge, feelings, and experiences, the ability of metacognition, is assumed to be a unique human capacity.

Metacognition has been defined as belonging to the key competences of the so-called 21st Century Skills. Researchers, teacher educators, and policy makers increasingly point to the importance of metacognition for success in learning from a lifelong perspective. In the language classroom, it is essential that language teachers are aware of their own knowledge and beliefs about languages, cultures, and language learning in order to support their students in their learning efforts. Similarly, language students benefit from a conscious approach to the language learning process (Haukas, Bjorke, and Dypedahl 2018). As the discussion in this chapter shows, the importance of metalinguistic awareness in the language learning process seems to increase along with a growth in intensity and amount of exposure to other languages. There is great necessity in multilingualism research for an in-depth study of metalinguistic awareness. That is especially the case for strategic processing in experienced, multilingual language learners, given the fact that research in SLA does not offer an accurate framework for this kind of study, and therefore cannot be applied to highly complex and dynamic systems, such as those of multilinguals. Consequently, the good language learner is an experienced language learner who is able to make use of her multilingual competencies as a strategic resource that helps to manage and direct further (or lifelong) language learning.

From a research perspective, this means that, only if we move away from a simplistic picture of language learning by taking the hypercomplexity of the multilingual mind into consideration, will we be able to make progress in understanding how language learning takes place (Aronin and Jessner 2015). To understand a multilingual person as somebody who has a different way of using and knowing her or his languages (in contrast to the native speakers of the respective languages) means that we acknowledge the cognitive opportunities that a life of multilingualism can offer, and profit from the benefits of contact with two or more languages.

6. References

Aronin, Larissa & Ulrike Jessner
2015 Understanding current multilingualism: What can the butterfly tell us? In Ulrike Jessner & Claire J. Kramsch (eds.), *The multilingual challenge: Cross-disciplinary perspectives*, 271–294. Berlin & Boston: De Gruyter Mouton.

Bialystok, Ellen
2001 Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics* 21. 169–181.

Bialystok, Ellen & Raluca Barac
2012 Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development* 83(2). 413–422.

Bialystok, Ellen, Fergus Craik, Raymond Klein & Mythili Viswanathan
2004 Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging* 19. 290–303.

Birdsong, David
1989 *Metalinguistic performance and interlinguistic competence*. Berlin: Springer.

Chaudron, Craig
1983 Research on metalinguistic judgments: A review of theory, methods and results. *Language Learning* 33(3). 345–377.

Cook, Vivian J.
1991 The poverty-of-the-stimulus argument and multi-competence. *Second Language Research* 7(2). 103–117.

Cook, Vivian J.
2002 Background to the L2 user. In Vivian Cook (ed.), *Portraits of the L2 user*, 1–28. Clevedon: Multilingual Matters.

Cots, Josep & Peter Garrett
2018 Language awareness: Opening up the field of study. In Peter Garrett & Josep M. Cots (eds.), *The Routledge handbook of language awareness*. London: Routledge. 1–19.

Council of Europe
2001 *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, U.K: Press Syndicate of the University of Cambridge.

Cummins, Jim
1991 Language learning and bilingualism. *Sophia Linguistica* 29. 1–194.

Cummins, Jim
2017 Teaching for transfer. In Stephen May, Ofelia Garcia & Angel Lin (eds.), *Encyclopedia of language and education, Vol 5., Bilingual programs*, 113–115. New York: Springer.

Dahm, Rebecca
2015 Developing cognitive strategies through pluralistic approaches. In Gessica De Angelis, Ulrike Jessner & Marijana Kresic (eds.), *Cross-linguistic influence and crosslinguistic interaction in multilingual language learning*, 43–70. London: Bloomsbury.

De Angelis, Gessica, Ulrike Jessner & Marijana Kresic (eds.)
2015 *Crosslinguistic influence and crosslinguistic interaction in multilingual language learning*. London: Bloomsbury Academic.

de Bot, Kees
2017 The future of the bilingual advantage. In Judit Navracsics & Simone Pfenninger (eds.), *Future research directions for applied linguistics*, 15–32. Bristol: Multilingual Matters.

Ellis, Nick C.
2011 Implicit and explicit SLA and their interface. In Cristina Sanz & Ronald P. Leow (eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA & bilingualism*, 35–47. Washington: Georgetown University Press.

Flynn, Suzanne, Claire Foley & Inna Vinnitskaya
2004 The cumulative-enhancement model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition of relative clauses. *International Journal of Multilingualism* 1(1). 3–16.
Furlong, Aine
2009 The relation of plurilingualism/culturalism to creativity: A matter of perception. *International Journal of Multilingualism* 6(4). 343–368.
Gnutzmann, Claus
1997 Language awareness: Progress in language learning and language education, or reformulation of old ideas? *Language Awareness* 6(2/3). 65–74.
Gombert, Jean-Emile
1992 *Metalinguistic development*. New York: Harvester Wheatsheaf.
Hamers, Josiane & Michel Blanc
1989 *Bilinguality and bilingualism*. Cambridge: Cambridge University Press.
Haukas, Asta, Camilla Bjorke & Magne Dypedahl (eds.)
2018 *Metacognition in language learning*. London: Routledge (OA).
Hawkins, Eric
1984 *Awareness of language: An introduction*. Cambridge: Cambridge University Press.
Herdina, Philip & Ulrike Jessner
2002 *A Dynamic Model of Multilingualism: Changing the psycholinguistic perspective*. Clevedon: Multilingual Matters.
Hofer, Barbara & Ulrike Jessner
2016 Multilingualism at the primary level in South Tyrol: How does multilingual education affect young learners' metalinguistic awareness and proficiency in L1, L2 and L3? *The Language Learning Journal*. http://dx.doi.org/10.1080/09571736.2016.1195865 (accessed 02 July 2017).
James, Carl
1996 A cross-linguistic approach to language awareness. *Language Awareness* 5(3&4). 138–148.
Jessner, Ulrike
2006 *Linguistic awareness in multilinguals: English as a third language*. Edinburgh: Edinburgh University Press.
Jessner, Ulrike
2008 A DST-model of multilingualism and the role of metalinguistic awareness. *Modern Language Journal* 92(2). 270–283.
Jessner, Ulrike
2016 Language awareness in multilinguals: Theoretical trends. In Stephen May, Jasone Cenoz & Durk Gorter (eds.), *Encyclopedia of language and education*, *Vol 6.: Knowledge about language*, 357–370. New York: Springer.
Jessner, Ulrike
2017 Multicompetence approaches to language proficiency development in multilingual education. In Stephen May, Ofelia Garcia & Angel Lin (eds.), *Encyclopedia of language and education*, *Vol 5.: Bilingual programs*, 161–173. New York: Springer.
Jessner, Ulrike & Valentina Török
2017 Strategies in multilingual learning: Opening new research avenues. In Judit Navracsics & Simone Pfenninger (eds.), *Future research directions for applied linguistics*, 192–214. Bristol: Multilingual Matters.
Jessner, Ulrike, Elisabeth Allgäuer-Hackl & Barbara Hofer
2016 Emerging multilingual awareness in educational contexts: From theory to practice. *The Canadian Modern Language Review / La revue canadienne des langues vivantes* 72(2). 157–182.

Jessner, Ulrike, Manon Megens & Stefanie Graus
2016 Crosslinguistic influence in Third Language Acquisition. In Rosa Alonso Alonso (ed.), *Crosslinguistic influence* in *Second Language Acquisition*, 193–214. Clevedon: Multilingual Matters.
Jessner, Ulrike, Valentina Török & Claudia Pellegrini
2018 The role of Latin in multilingual learners' strategies. *Language Education and Multilingualism* 1. 85–102.
Karmiloff-Smith, Annette
1992 *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
Kecskes, Istvan & Tünde Papp
2000 *Foreign language and mother tongue*. Mahwah, NJ: Lawrence Erlbaum.
Kemp, Charlotte
2001 *Metalinguistic awareness in multilinguals: Implicit and explicit grammatical awareness and its relationship with language experience and language attainment*. Edinburgh: University of Edinburgh dissertation.
Kharkhurin, Anatoliy V.
2012 *Multilingualism and creativity.* Bristol: Multilingual Matters.
Leow, Ronald P.
2015 Implicit learning in SLA: Of processes and products. In Patrick Rebuschat (ed.), *Implicit and explicit learning of languages*, 47–66. Amsterdam & Philadelphia: John Benjamins.
Leow, Ronald P., Ellen Johnson & German Zarate-Sandez
2011 Getting a grip of the slippery constract of awareness: Towards a finer-grained methodological perspective. In Cristina Sanz & Ronald P. Leow (eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism*, 61–72. Washington: Georgetown University Press.
Mitits, Lydia
2015 *Language learning strategies and multilingualism: Monolingual EFL and multilingual EFL/L2 Greek learners in Greek secondary education*. Kavala: Saita Publications.
Modirkhamene, Sima
2008 Metalinguistic awareness and bilingual vs. monolingual EFL learners: Evidence from a diagonal bilingual context. *The Journal of Applied Linguistics* 1(1). 66–102.
Mohanty, Ajit K.
1994 *Bilingualism in a multilingual society: Psychosocial and pedagogical implications*. Mysore: Central Institute of Indian Languages.
Montanari, Simona
2009 Pragmatic differentiation in early trilingual development. *Journal of Child Language* 36(3). 597–627.
Moore, Danièle
2006 *Plurilinguisme et Ecole.* Paris: Didier.
Nation, Robert & Barry McLaughlin
1986 Experts and novices: An information-processing approach to the 'good language learner' problem. *Applied Psycholinguistics* 7. 51–56.
Nayak, Nandini, Nina Hansen, Nancy Krueger & Barry McLaughlin
1990 Language-learning strategies in monolingual and multilingual adults. *Language Learning* 40(2). 221–244.
Pinto, Maria A. (ed.)
2015 The "MATEL" (Metalinguistic awareness tests in European languages) project: Validations and translations. [Special issue]. *Rivista di Psicolinguistica Applicata: Journal of Applied Psycholinguistics* 15(2).

Pinto, Maria A., Renzo Titone & Francesca Trusso
1999 *Metalinguistic awareness: Theory, development and measurement instruments*. Pisa & Rome: Istituti Editoriali e Poligrafici Internazionali.
Psaltou-Joycey, Angeliki & Zoe Kantaridou
2009 Foreign language learning strategy profiles of university students in Greece. *Journal of Applied Linguistics* 25. 107–127.
Safont-Jordà, Maria Pilar
2013 Early stages of trilingual pragmatic development: A longitudinal study of requests in Catalan, Spanish and English. *Journal of Pragmatics* 59. 68–80.
Simard, Daphnée & Xavier Guitérrez
2018 The study of metalinguistic constructs in second language acquisition research. In Peter Garrett & Josep M. Cots (eds.), *The Routledge handbook of language awareness*, 205–221. London: Routledge.
Troncy, Christel (ed.)
2014 *Didactique du plurilinguisme. Approches plurielles des langues et des cultures: Autour de Michel Candelier*. Rennes: Presses Universitaires de Rennes.
Young, Andrea & Christine Hélot
2003 Language awareness and/or language learning in French primary schools today. *Language Awareness* 12(3&4), 234–246.
van Essen, Arthur
2008 Language awareness and knowledge about language: A historical overview. In Jasone Cenoz & Nancy H. Hornberger (eds.), *Encyclopedia of language and education (Knowledge about language, 2nd edn, Vol. 6)*, 3–14. New York: Springer.
Vygotsky, Lew
1986 [1934] *Thought and Language*. Cambridge, MA: MIT Press. [revised, translated, and edited by Alex Kozulin]

Ulrike Jessner, Innsbruck (Austria) and Veszprem (Hungary)

20. Language attitudes

1. General introduction
2. The early history of research on language attitudes
3. Methodological approaches to research on language attitudes
4. More recent developments in language attitudes research
5. Outlook
6. References

1. General introduction

Language attitudes develop in the context of group contact, often between groups using different first languages. The next section below provides a brief overview of early work on language attitudes. Then, some current issues in this field are outlined. The last section provides a brief outlook of directions in which language attitudes research ap-

https://doi.org/10.1515/9783110435351-020

pears to be moving. Examples of language attitudes research conducted in language contact settings are highlighted.

2. The early history of research on language attitudes

The study of attitudes within social psychology has informed the field of language attitudes. In order to understand the latter, the controversies that have characterized the former are of relevance, e.g. the overall nature of attitudes (whether observable and directly related to behavior, or not; unicomponential or multicomponential), or the relationship between attitudes and stereotypes, ideologies, perceptions, beliefs, etc. Most theorists conceptualize attitudes as composed of three interrelated components: affective, cognitive, and conative (action). Oppenheim (1966: 106–111) still provides a satisfying account: "[a]ttitudes are reinforced by beliefs (the cognitive component) and often attract strong feelings (the emotional component) that will lead to particular forms of behavior".

An important role in attitude studies in general has been played by Social Identity Theory (SIT) (Tajfel 1978; Tajfel and Turner 1979), which has been concerned from the beginning with group (and thus language) contact (Tajfel and Turner 2004). In terms of SIT, social identity begins with social categorization (Hogg and Abrams 1988: 24–5): people are perceived in terms of whether or not they belong to the same social group as oneself. Generally, for members of a group, it is necessary to ensure positive social identity, and to accentuate features which reflect positively on in-group (and often negatively on outgroup) members. This ensures positive distinctiveness for the in-group. The success of this operation is, however, often determined by the power relations existing between groups in contact (Hogg and Abrams 1988: 26–7). Powerful groups are able to disseminate their own ideologies and stereotypes, and entrench them as common sense, while minority groups are often under pressure to adopt these. Minority-group members thus often have a negative social identity, and are motivated to change their identity via strategies such as assimilation into the more powerful group.

Language is often an index of group membership, as explored by modern sociolinguistics (Eckert 1989; Labov 1963). The language varieties of groups in contact often become the focus of social comparison designed to create positive distinctiveness. Most social groups are indexed linguistically, but it is possible to isolate a number of groups for whom language often plays a central symbolic role; these include social classes and ethnic groups, the latter often speaking different languages and initiating language contact.

The complex relationships between language, ethnicity, and power have been elaborated by the 'theory of language in ethnic-group relations', as first developed by Giles, Bourhis, and Taylor (1977), combining as it does SIT, speech accommodation theory, and the notion of 'ethnolinguistic vitality', which is what "makes a group likely to behave as a distinct and active collective entity in intergroup situations" (Giles, Bourhis, and Taylor 1977: 308), i.e. in language contact settings. Theoretical discussion on the relationship between language and ethnicity is often part of more general treatments of language maintenance and shift (Edwards 1985; Paulston 1994). Assuming attitudes are related to behavior, active intervention (e.g. language-policy initiatives) aimed at changing ethnolinguistic vitality, can have an effect on language attitudes, and thus on (de)accelerating language shift (Eastman 1983: 58).

The relationship between social class and language has been a focus of sociolinguistics since its inception (Labov 1966). The study of language attitudes is a core aspect of the sociolinguistic project, not least because the definition of 'speech community' as a social group depends on shared language attitudes. Naturally, class often overlaps with other social variables, e.g. ethnicity. In the case of a caste system, where each class is often associated with a different ethnic group, the notion of diglossia (and concomitant attitudes) has relevance. Of interest are increasingly common cases in which middle-class identity and a particular language overlap, but where the language in question has to a large degree lost its ethnolinguistic associations. To an increasing degree, this is the case with English as an international language, which is perceived as a language of social mobility, but is often (depending on where in the world the user lives) stripped of its colonial, racial, and/or Anglo-American associations (see, for example, Kamwangamalu and Tovares 2016). Of relevance here is the distinction proposed by Gardner and Lambert (1972) between instrumental and integrative orientations towards language learning, with English often being learned for instrumental, pragmatic reasons alone.

3. Methodological approaches to research on language attitudes

With regard to methodology, three approaches dominate the field of language attitudes. As described by Garrett (2010: 37), these are: "the analysis of the societal treatment of language varieties […] direct measures and indirect measures" (see also Agheyisi and Fishman 1970: 141–150; Giles, Hewstone, and Ball 1983: 82–86; Ryan, Giles, and Hewstone 1982: 6–8; Ryan, Giles, and Hewstone 1988).

Analyses of societal treatment (e.g. Bishop, Coupland, and Garrett 2005; Cheshire and Moser 1994; Fishman 1966; Rickford and Traugott 1985) subsume a range of qualitative techniques "which do not involve explicitly asking respondents for their views or reactions" (Ryan, Giles, and Hewstone 1988: 1068). Examples include "demographic and census analyses; analyses of government and educational language policies; analyses of literature, government and business documents, newspapers and broadcasting media" (Ryan, Giles, and Hewstone 1988: 1068), as well as other methodologies, e.g. participant-observation and case-studies. Such studies offer a broad overview of the socio-historical roots of language attitudes (Bekker 2003) and highlight the dynamic nature of language attitudes, providing "some insight into the emergence of values over time" (Giles, Hewstone, and Ball 1983: 83).

Direct measurement mainly involves interviews and questionnaires. While "some scholars dismiss the use of direct methods […] completely" (Smit 1996: 45), the main objection being their inability to "reveal non-consciously held attitudes" (Giles, Hewstone, and Ball 1983), in practice they have been employed in research on a range of related topics, including language preference, the prediction of second-language learning outcomes, attitudes towards language-policy, and bilingualism and bilingual education (Garrett 2010: 159–178; Ryan, Giles, and Hewstone 1982: 7; Ryan, Giles, and Hewstone 1988: 1069–1070). With respect to closed questionnaires in particular, researchers benefit from using instruments constructed for attitude measurement more generally; thus the Thurstone, Guttman, and Likert Scales have all been more or less successfully used in language attitudes research (Smit 1996: 46). Examples of the use of these

methods include: Fishman (1966), Cohen (1974), Hyrkestedt and Kalaja (1998), BBC (2005), and Garrett, Bishop, and Coupland (2009).

Turning to indirect methods, the Matched-Guise Technique (MGT), as first introduced by Lambert et al. (1960), is the only method specifically designed for language attitudes research. Its basic form involves tape-recording bidialectal speakers reading a text in usually two (native or non-native) dialects (e.g. Edwards 1985: 150; Giles, Hewstone, and Ball 1983: 84; Giles et al. 1988). The two versions (with distractors interspersed in-between) are played to respondents, who are asked to rate 'each speaker' on a number of personality traits using semantic differential scales (Osgood, Suci, and Tannenbaum 1957). The respondents are (ideally) unaware that the two varieties are from the same speaker and the idea is that "whatever evaluations are made of the speaker must be prompted mainly by the judge's general reaction to the speakers of that particular language, rather than by his reaction to the specific speaker in the experimental situation" (Agheyisi and Fishman 1970: 146). The main advantage of this method is that it negates the problem of respondents providing expected answers. It is also meant to tap into the affective, implicit side of language attitudes (Smit 1996: 47). A number of variants have arisen, e.g. the 'verbal guise study' (Cooper 1974), in which the different 'guises' are produced by different (although similar) speakers. This variant was motivated by concerns "regarding the practical implausibility of finding a speaker who can convincingly produce authentic speech samples of the language varieties presented for evaluation" (McKenzie 2015: 33). Of methodological concern has also been the nature of the scales used (e.g. odd or even numbers; issues relating to labelling), as well as the nature and number of evaluative dimensions. While the use of factor-analysis by Zahn and Hopper (1985) led to the discovery that semantic scales generally "loaded into three […] factors […] 'superiority' […] 'attractiveness' […] and 'dynamism'" (Garrett 2010: 55), the two dimensions superiority/status and attractiveness have generally received the most attention in the literature.

4. More recent developments in language attitudes research

Of particular relevance to the topic of language contact has been recent work done on the role of social categorization, and how differences in categorization accuracy activate different stereotypes, and thus different language attitudes (McKenzie 2008; Yook and Lindemann 2013). Thus Dragojevic, Berglund, and Blauvelt (2017: 2) confirm in their study of different ethnolinguistic groups in Georgia that "intergroup variation in categorization accuracy may produce intergroup variation in language attitudes" and also how "language attitudes may be far more variable and dynamic than existing research suggests […] the *same* speaker may garner markedly different evaluations in different situations, simply because they may be categorized differently" (Dragojevic, Berglund, and Blauvelt 2017: 16). These same authors emphasize that, in real-life contexts, categorization and resultant attitudes depend not solely on linguistic features, but on other cues, such as "speakers' skin color, dress, and so forth" (Dragojevic, Berglund, and Blauvelt 2017: 17). Further research has explored a range of other factors which complicate (as intervening variables) the link between a linguistic variable and its indexical meaning, e.g. the degree to which respondents can "recognize a speaker's *degree* of accentedness"

(Mirshahidi 2017: 156) or "the ease or difficulty associated with processing speech in [the] variety" (Dragojevic et al. 2017: 401), whether due to purely environmental factors (Dragojevic and Giles 2016) or to "disruptions in fluency due to accent *itself*" (Dragojevic et al. 2017: 386).

Of further relevance to language contact is research conducted on language attitudes within an L2- and/or L3-learning context in relation to a variety of different settings, e.g. Spain (Lasagabaster 2017), Japan (McEown, Sawaki, and Harada 2017), and Hungary (Csizér and Lukács 2010). Such research has been characterized by a number of themes, including different theories of language learning motivation (e.g. Teimouri 2017), the link between language attitudes and language practices, and the effect of English as a global lingua franca on attitudes towards learning other languages, as well as language attitude change and the different trajectories that different individuals/groups take in terms of such change (Ianos et al. 2017; Ubalde, Alarcón, and Lapresta 2017). Much of this research, particularly that which was conducted in Spain, has focused on the language attitudes of immigrant communities "in areas where two or more languages are in contact" (Ianos et al. 2017: 331).

Continuing with language contact contexts that arise via immigration and other sources of group-contact (e.g. multilingual international corporations), the indirect methods of measurement discussed at the end of the last section are still relevant in terms of the discovery of "underlying prejudices […] which might not, or cannot, be otherwise revealed through the use of explicit measures" (McKenzie 2015: 47). Probably the most important recent development on a methodological level with respect to indirect methods has been the use of the Implicit Association Test (IAT), developed within the field of social psychology (Greenwald, McGhee, and Schwartz 1998), and very recently applied and adapted to the field of language-attitude research (Álvarez-Mosquera and Marín-Gutiérrez 2017; Babel 2010; Campbell-Kibler 2012; Pantos and Perkins 2013; Redinger 2010; Rosseel, Speelman, and Geeraerts 2015). The IAT can be characterized in general as "an effective method to assess the strength of the association between concepts and evaluations by means of participants' reaction times toward given stimuli" (Álvarez-Mosquera and Marín-Gutiérrez 2017: 1–2). With respect to its application to language attitudes research, the relevant association is between linguistic forms and attitudes towards those forms; the stimuli are often of both an audio and textual nature. The IAT provides the advantages of an indirect method, but also avoids some of the problems inherent in the application of the MGT. It has been particularly useful in tapping into the difference between implicitly and explicitly held attitudes. So for example, Pantos (2015) found that "results revealed a pro-US accent bias on the implicit measure, but a pro-foreign accent bias on explicit measures".

One modification in the application of the MGT has been the manipulation of speech variables (as opposed to the use of whole dialects) via the splicing, grafting, and general manipulation of sounds using modern acoustic software (Campbell-Kibler 2007; Levon 2006). Thus, by way of example, Bekker and Levon (2017) recorded balanced Afrikaans-English bilinguals reading the same text in both languages, once with a 'standard' alveolar /s/ and once with a fronted /s/. The instances of fronted /s/ were then spliced and grafted into the non-fronted-/s/ recordings. This application of the matched, or verbal-guise, technique is of particular relevance to sociolinguistic investigations on the indexical value of particularly phonetic variants.

The field of social psychology has, naturally, undergone change since the inception of work in language attitudes research, not least of which is the relatively recent turn to a social-constructivist, discourse-oriented approach towards the study of attitudes, a canonical text in this regard being Potter and Wetherell (1987). This relatively new approach has had an effect on the study of language attitudes, with a particular emphasis on challenging their perceived stability and durability across individuals, contexts, and time (Garrett 2010: 29–30). An interesting example of this is a recent study by Cortès-Colomé, Barrieras, and Comellas (2016), which shows how the attitudes of immigrants towards their own language changes (in positive ways) as a result of increased exposure to a successful minority language, in this case Catalan. Of interest here, too, are a number of theoretical interventions based less on the Anglo-American social psychology tradition, and more on continental-European traditions with a greater focus on conflict and class-struggle, often with a (neo-)Marxist flavor. A pertinent example here would be Bourdieu's (1991) notions of *linguistic market* and *linguistic capital*, the latter being of particular relevance to multilingual contexts (such as are common in Africa), in which the linguistic market often leads to a "certain amount of linguistic schizophrenia […] in which Africans admire educated English, especially the learned style, and tend to cultivate it, while avoiding speaking "too well" or with a standard accent" (Schmied 1991: 173). Language attitudes (and their link to issues of identity in particular) play a key role in attempts to explain the evolution of languages in (post-)colonial contexts (Schneider 2007).

More generally, studies focused on the broad relationship between language(s) and ideology (e.g. Lippi-Green 1997; Tollefson and Yamagami 2012) have increased our scope of understanding when it comes to language attitudes, a good recent example of this being Atkinson and Kelly-Holmes (2016: 199), whose exploration "of deep ambivalence towards the use of the Irish language, as well as complex ideological positions" helps explain the discrepancy between positive language attitudes toward, but low levels of practical use of, Irish on the one hand, while on the other hand assigning greater levels of authenticity to other (immigrant) minority languages with which Irish is in contact, e.g. Polish. With reference to another kind of language contact situation (i.e. not involving immigration), Nguyen and Hamid (2018: 343) provide a similarly recent example of the exploration of positive language attitudes/ideologies towards bilingualism per se, i.e. "bilingualism as a "resource" that [can] help […] to maintain […] ethnic features and join the mainstream flow simultaneously". Sociolinguistics, whether of the classical (Labov 1963, 1966) or more modern (Eckert 1989; Levon and Buchstaller 2015; Rampton 2005) kind, provides similar insight into the indexical and semiotic value of linguistic varieties and features, as well as growing insight into the link between such indexical value and language change (Kristiansen 2015). Thus, Labov's (1963) conclusions regarding the indexical associations of centralized /ay/ and /au/ on Martha's Vineyard quite clearly provide us with insight into the language attitudes of the relevant speakers, and how these attitudes are instrumental in bringing about the relevant changes. A related, recently developed direct method of language attitudes research is folk linguistics (Garrett 2010: 179–200), which, as the name suggests, elicits the perceptions, ideologies, etc. of non-linguists on topics that linguists generally reserve as their own. A core component is perceptual dialectology, often employing map tasks (Niedzielski and Preston 2000; Preston 1989).

5. Outlook

The discursive turn in social psychology (Potter and Wetherell 1987) and so-called third-wave sociolinguistics (Eckert 2012) has meant that the notion of fixed associations and attitudes towards language has come under increased interrogation, while greater emphasis is being placed on the contextualized, stylistic use of language to create social meaning (Rampton 2005). The indexical link between language and society − the associations and attitudes employed and expressed by those involved − is, in this account, a fluid one. Of relevance here is work in (critical) discourse analysis (Fairclough [1989] 2014), with its focus on how linguistic choice actuates ideology. Our responses to (and attitudes toward) a particular text are driven by an (often implicit) understanding of underlying ideologies, and the associations and attitudes cued by the linguistic features of the (written or spoken) text are powerfully dependent on contextual features, such as audience and situational context (Bell 1984). Such work is likely to play a central role in future research into language attitudes (see for example Purschke [2015] for a new theoretical framework that incorporates this emphasis), while this overall move towards viewing language attitudes in the light of individual, highly-contextualized responses to (mostly) spoken texts opens the door to this area of linguistics being influenced more and more by theories of language which emphasize context-specific factors, and even by concepts prevalent in literary theory. So, for example, Bakhtin's (1973, 1981) work, with its emphasis on the unique utterance, polyphony, and heteroglossia, would be of particular relevance to research focused on the use of indexical meaning to create irony, parody, and other subtle indexical/attitudinal effects, e.g. work on the use of phonetic variables to cue indexical meaning in modern forms of media, such as music videos (Williams and Stroud 2014). Another important new field which assists in understanding the ideological value of language varieties in their real-world context is that of linguistic landscapes (Shohamy and Gorter 2009).

In the previous section, Dragojevic, Berglund, and Blauvelt (2017) were shown to emphasize the role played, in real-life contexts, of non-linguistic cues on social categorization (and indirectly on language attitudes). Future research will have to take such real-life factors into account, and in particular, to examine "how multiple cues to speakers' social group membership(s) − which may complement *or* contradict one another − influence the categorization process" (Dragojevic, Berglund, and Blauvelt 2017: 18). Other intervening variables, such as personality characteristics, also warrant further investigation (Hansen and Dovidio 2016). More generally, there is still an important role for research using indirect measures of language attitudes and related prejudices, particularly in language contact contexts, such as multinational corporations, where individuals are unlikely to report their prejudices or "their feelings of social marginalization and/or experiences of perceived unfairness to researchers" (Woo and Giles 2017: 48).

6. References

Agheyisi, Rebecca & Joshua A. Fishman
1970 Language attitude studies: A brief survey of methodological approaches. *Anthropological Linguistics* 12(5). 137−157.

Álvarez-Mosquera, Pedro & Alejandro Marín-Gutiérrez
2017 Implicit language attitudes toward historically white accents in the South African context. *Journal of Language and Social Psychology* 37(2). 238–248.
Atkinson, David & Helen Kelly-Holmes
2016 Exploring language attitudes and ideologies in university students' discussions of Irish in a context of increasing language diversity. *Language and Intercultural Communication* 16(2). 199–215.
Babel, Molly
2010 Dialect convergence and convergence in New Zealand English. *Language in Society* 39(4). 437–456.
Bakhtin, Mikhail/Voloshinov, Valentin N.
1973 *Marxism and the philosophy of language*. London: Harvard University Press.
Bakhtin, Mikhail
1981 *The dialogic imagination*. Austin: University of Texas Press.
BBC
2005 Regional accents 'bad for trade'. http://news.bbc.co.uk/2/hi/uk_news/england/4566028.stm (accessed on 7 March 2017).
Bekker, Ian
2003 Using historical data to explain language attitudes: A South African case study. *AILA Review* 16. 62–77.
Bekker, Ian & Erez Levon
2017 The embedded indexical value of /s/-fronting in Afrikaans and South African English. *Linguistics* 55(5). 1109–1139.
Bell, Allan
1984 Language style as audience design. *Language in Society* 13(2). 145–204.
Bishop, Hywel, Nikolas Coupland & Peter Garrett
2005 Globalisation, advertising and language choice: Shifting values for Welsh and Welshness. *Multilingua* 24. 343–78.
Bourdieu, Pierre
1991 *Language and symbolic power*. Cambridge, MA: Harvard University Press.
Campbell-Kibler, Kathryn
2007 Accent, (ING), and the social logic of listener perceptions. *American Speech* 82(1). 32–64.
Campbell-Kibler, Kathryn
2012 The Implicit Association Test and sociolinguistic meaning. *Lingua* 122. 753–763.
Cheshire, Jenny & Lise-Marie Moser
1994 English as a cultural symbol: The case of advertisements in French-speaking Switzerland. *Journal of Multilingual and Multicultural Development* 15(6). 451–69.
Cohen, Andrew D.
1974 Mexican-American evaluational judgements about language varieties. *International Journal of the Sociology of Language* 3. 33–51.
Cooper, Robert L. (ed.)
1974 The study of language attitudes. [Special issue]. *International Journal of the Sociology of Language* 3.
Cortès-Colomé, Montserrat, Mònica Barrieras & Pere Comellas
2016 Changes in immigrant individuals' language attitudes through contact with Catalan: The mirror effect. *Language Awareness* 25(4). 272–289.
Csizér, Kata & Gabriella Lukács
2010 The comparative analysis of motivation, attitudes and selves: The case of English and German in Hungary. *System* 38(1). 1–13.

Dragojevic, Marko & Howard Giles
2016 I don't like you because you're hard to understand: The role of processing fluency in the language attitudes process. *Human Communication Research* 42. 396–420.
Dragojevic, Marko, Christofer Berglund & Timothy K. Blauvelt
2017 Figuring out who's who: The role of social categorization in the language attitudes process. *Journal of Language and Social Psychology* 37(1). 28–50.
Dragojevic, Marko, Howard Giles, Anna-Carrie Beck & Nicholas T. Tatum
2017 The fluency principle: Why foreign accent strength negatively biases language attitudes. *Communication Monographs* 84(3). 385–405.
Eastman, Carol M.
1983 *Language planning: An introduction*. San Francisco: Chandler & Sharp.
Eckert, Penelope
1989 *Jocks and burnouts: Social categories and identity in the high school*. New York & London: Teachers College Press.
Eckert, Penelope
2012 Three waves of variation study: The emergence of meaning in the study of variation. *Annual Review of Anthropology* 41. 87–100.
Edwards, John
1985 *Language, society and identity.* Oxford: Basil Blackwell.
Fairclough, Norman
2014 [1989] *Language and power*, 3rd edn. London: Routledge.
Fishman, Joshua A. (ed.)
1966 *Language loyalty in the United States: The maintenance and perpetuation of non-English mother-tongues by American ethnic and religious groups*. The Hague: Mouton.
Gardner, Robert C. & Wallace E. Lambert
1972 *Attitudes and motivation in second-language learning*. Rowley, MA: Newbury House.
Garrett, Peter
2010 *Attitudes to language*. Cambridge: Cambridge University Press.
Garrett, Peter, Hywel Bishop & Nikolas Coupland
2009 Diasporic ethnolinguistic subjectivities: Patagonia, North America and Wales. *International Journal of the Sociology of Language* 195. 173–99.
Giles, Howard, Miles Hewstone & Peter Ball
1983 Language attitudes in multilingual settings: Prologue with priorities. *Journal of Multilingual and Multicultural Development* 4(2/3). 81–96.
Giles, Howard, Miles Hewstone, Ellen B. Ryan & Patricia Johnson
1988 Research on language attitudes. In Ulrich Ammon, Norbert Dittmar & Klaus J. Mattheier (eds.), *Sociolinguistics: An international handbook of the science of language*, vol. 2, 585–597. Berlin: De Gruyter.
Giles, Howard, Richard Y. Bourhis & Donald M. Taylor
1977 Towards a theory of language in ethnic-group relations. In Howard Giles (ed.), *Language, ethnicity and intergroup relations*, 317–348. London: Academic Press.
Greenwald, Anthony G., Debbie E. McGhee & Jordan L. K. Schwartz
1998 Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology* 74(6). 1464–1480.
Hansen, Karolina & John F. Dovidio
2016 Social dominance orientation, nonnative accents, and hiring recommendations. *Cultural Diversity and Ethnic Minority Psychology* 22(4). 544–551.
Hogg, Michael A. & Dominic Abrams
1988 *Social identifications: A social psychology of intergroup relations and group processes*. London: Routledge.

Hyrkestedt, Irene & Paula Kalaja
1998 Attitudes towards English and its functions in Finland: A discourse-analytic study. *World Englishes* 17. 345–57.
Ianos, Maria-Adelina, Ángel Huguet, Judit Janés & Cecilio Lapresta
2017 Can language attitudes be improved? A longitudinal study of immigrant students in Catalonia (Spain). *International Journal of Bilingual Education and Bilingualism* 29(3). 331–345.
Kamwangamalu, Nkonko & Alla Tovares
2016 English in language ideologies, attitudes, and educational practices in Kenya and South Africa. *World Englishes* 35(3). 421–439.
Kristiansen, Tore
2015 The primary relevance of subconsciously offered attitudes: Focusing the language ideological aspect of sociolinguistic change. In Alexei Prikhodkine & Dennis R. Preston (eds.), *Responses to language varieties: Variability, processes and outcomes*, 87–116. Amsterdam: John Benjamins.
Labov, William
1963 The social motivation of a sound change. *Word* 19(3). 273–309.
Labov, William
1966 *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
Lambert, Wallace E., Robert Hodgson, Robert C. Gardner & Steven Fillenbaum
1960 Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60. 44–51.
Lasagabaster, David
2017 Language learning motivation and language attitudes in multilingual Spain from an international perspective. *The Modern Language Journal* 101(3). 583–596.
Levon, Erez
2006 Hearing "gay": Prosody, interpretation and the affective judgements of men's speech. *American Speech* 81(1). 56–78.
Levon, Erez & Isabelle Buchstaller
2015 Perception, cognition and linguistic structure: The effect of linguistic modularity and cognitive style on sociolinguistic processing. *Language Variation and Change* 27(3). 319–348.
Lippi-Green, Rosina
1997 *English with an accent: Language, ideology, and discrimination in the United States*. London: Routledge.
McEown, Maya S., Yasuyo Sawaki & Tetsuo Harada
2017 Foreign language learning motivation in the Japanese context: Social and political influences on self. *The Modern Language Journal* 101(3). 533–547.
McKenzie, Robert M.
2008 The role of variety recognition in Japanese university students' attitudes towards English speech varieties. *Journal of Multilingual and Multicultural Development* 29(2). 139–153.
McKenzie, Robert M.
2015 UK university students' folk perceptions of spoken variation in English: The role of explicit and implicit attitudes. *International Journal of the Sociology of Language* 236. 31–53.
Mirshahidi, Shahriar
2017 I find you attractive but I don't trust you: The case of language attitudes in Iran. *Journal of Multilingual and Multicultural Development* 38(2). 146–159.
Niedzielski, Nancy A. & Dennis R. Preston
2000 *Folk linguistics*. New York: Mouton.

Nguyen, Trang T. T. & M. Obaidul Hamid
2018 Bilingualism as a resource: Language attitudes of Vietnamese ethnic minority students. *Current Issues in Language Planning* 19(4). 343–362.
Oppenheim, Abraham N.
1966 *Questionnaire design and attitude measurement*. London: Heinemann.
Osgood, Charles E., George J. Suci & Percy H. Tannenbaum
1957 *The measurement of meaning*. Urbana: University of Illinois Press.
Pantos, Andrew J.
2015 Applying the Implicit Association Test to language attitudes research. In Alexei Prikhodkine & Dennis R. Preston (eds.), *Responses to language varieties: Variability, processes and outcomes*, 117–136. Amsterdam: John Benjamins.
Pantos, Andrew J. & Andrew W. Perkins
2013 Measuring implicit and explicit attitudes towards foreign-accented speech. *Journal of Language and Social Psychology* 32(1). 3–20.
Paulston, Christina B.
1994 *Linguistic minorities in multilingual settings*. Amsterdam: John Benjamins.
Potter, Jonathan & Margaret Wetherell
1987 *Discourse and social psychology: Beyond attitudes and behavior*. London: Sage.
Preston, Dennis R.
1989 *Perceptual dialectology: Nonlinguists' views of areal linguistics*. Dordrecht: Foris.
Purschke, Christoph
2015 REACT: A constructivist theoretic framework for attitudes. In Alexei Prikhodkine & Dennis R. Preston (eds.), *Responses to language varieties: Variability, processes and outcomes*, 37–53. Amsterdam: John Benjamins.
Rampton, Ben
2005 *Crossing: Language and ethnicity among adolescents*. Manchester, UK & Northampton, MA: St Jerome Publishing.
Redinger, Daniel
2010 *Language attitudes and code-switching behavior in a multilingual educational context: The case of Luxembourg*. York: University of York dissertation.
Rickford, John R. & Elizabeth C. Traugott
1985 Symbol of powerlessness and degeneracy, or symbol of solidarity and truth? Paradoxical attitudes towards pidgins and creoles. In Sidney Greenbaum (ed.), *The English language today*, 252–261. Oxford: Pergamon.
Rosseel, Laura, Dirk Speelman & Dirk Geeraerts
2015 Can social psychological attitude measures be used to study language attitudes? A case study exploring the Personalized Implicit Association Test. In Johannes Wahle, Marisa Köllner, Harald Baayen, Gerhard Jäger & Tineke Baayen-Oudshoorn (eds.), *Proceedings of the 6th conference on quantitative investigations in theoretical linguistics*, http://dx.doi.org/10.15496/publikation-8642. Tübingen: University of Tübingen (accessed on 25 January 2018).
Ryan, Ellen B., Howard Giles & Miles Hewstone
1982 An integrative perspective for the study of attitudes towards language variation. In Ellen B. Ryan & Howard Giles (eds.), *Attitudes towards language variation: Social and applied contexts*, 1–19. London: Edwin Arnold.
Ryan, Ellen B., Howard Giles & Miles Hewstone
1988 The measurement of language attitudes. In Ulrich Ammon, Norbert Dittmar & Klaus J. Mattheier (eds.), *Sociolinguistics: An international handbook of the science of language*, vol. 2, 1068–1081. Berlin: De Gruyter.
Schmied, Josef J.
1991 *English in Africa: An introduction*. London: Longman.

Schneider, Edgar
2007 *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
Shohamy, Elana & Durk Gorter (eds.)
2009 *Linguistic landscape: Expanding the scenery*. London: Routledge.
Smit, Ute
1996 *A new English for a New South Africa: Language attitudes, language planning and education*. Vienna: Braumüller.
Tajfel, Henri
1978 The achievement of inter-group differentiation. In Henri Tajfel (ed.), *Differentiation between social groups*, 77−100. London: Academic Press.
Tajfel, Henri & John C. Turner
1979 An integrative theory of inter-group conflict. In William G. Austin & Stephen Worchel (eds.), *The social psychology of inter-group relations*, 33−47. Monterey, CA: Brooks/Cole.
Tajfel, Henri & John C. Turner
2004 The social identity theory of intergroup behavior. In John T. Jost & Jim Sidanius (eds.), *Political psychology: Key readings*, 276−293. New York: Psychology Press.
Teimouri, Yasser
2017 L2 selves, emotions and motivated behaviors. *Studies in Second Language Acquisition* 39(4). 681−709.
Tollefson, James W. & Mai Yamagami
2012 Language ideology in a language classroom. *The Encyclopaedia of Applied Linguistics*. http://onlinelibrary.wiley.com/doi/10.1002/9781405198431.wbeal0625/full (accessed on 26 January 2018).
Ubalde, Josep, Amado Alarcón & Cecilio Lapresta
2017 Evolution and determinants of language attitudes among Catalan adolescents. *International Journal of Intercultural Relations* 60. 92−103.
Williams, Quentin E. & Christopher Stroud
2014 Battling the race: Stylizing language and coproducing whiteness and colouredness in a freestyle rap performance. *Journal of Linguistic Anthropology* 24(3). 277−93.
Woo, DaJung & Howard Giles
2017 Language attitudes and intergroup dynamics in multilingual organizations. *International Journal of Cross-cultural Management* 17(1). 39−52.
Yook, Cheongmin & Stephanie Lindemann
2013 The role of speaker identification in Korean university students' attitudes towards five varieties of English. *Journal of Multilingual and Multicultural Development* 34. 279−296.
Zahn, Christopher & Robert Hopper
1985 Measuring language attitudes: The speech evaluation instrument. *Journal of Language and Social Psychology* 4. 113−23.

Ian Bekker, Potchefstroom (South Africa)

21. Gender and language contact: how gender is/isn't marked in language contact

1. Introduction
2. Evidence of gender effects in language contact
3. Weak evidence for gender stratification in communities
4. Speculations on the gap
5. References

1. Introduction

Stereotypes about women and language abound. Girls are better at learning languages than boys. Women's talk is trivial, mean-spirited, and indecisive, and there is too much of it. Their intonation is irritating. Their vocabulary, e.g. the frequent use of *just* or *like*, shows expressive and/or cognitive deficiencies. They do not express themselves assertively enough in business or the professions. They don't get jokes. Cross-linguistic surveys (e.g. Hellinger and Bußmann 2001) show that such stereotypes are widespread and typically express negative perceptions. Like perhaps all language ideologies, the attitudes people express about *how* women talk are lightly-veiled attitudes about women themselves (see Cameron 2014); expressing negative evaluations of women's talk allows people to be negative about women, without being held directly accountable for sexism or gender bias. Paradoxically, as studies of language variation have shown repeatedly, the linguistic features women are negatively stereotyped for today may prove to be the innovations which the rest of the speech community is likely to follow in due course.

Notably, there are very few stereotypes − positive or negative − about how women or men behave in situations of language contact. Considering that most people in the world are in situations of (stable or changing) language contact (Nagy and Meyerhoff 2008), and considering that there are gender stereotypes about a wide range of linguistic behaviors and about language use in numerous social domains, an absence of social commentary about gender and language contact invites closer inspection.

In this chapter, we review literature on gender and language contact. We find ample evidence that individuals agentively draw on different languages or linguistic features when adopting particular (gendered) stances or attitudes in relation to others. We also find that communities often perceive the use of a particular language or linguistic feature to be 'feminine' or 'masculine'. Yet these qualitative findings are not matched by quantitative ones. It proves much harder to find evidence that gender is a significant predictor of variation: gender effects are very weak, inconsistent in the direction of their effect, or simply not present at all in variationist studies of language contact.

This is curious: gender emerges as a social constraint time and again in other studies of language variation and change (Labov 1994, 2001, 2010), and gender differences emerge in studies of language variation not only in Western, monolingual, urban speech communities (Eckert 1989; Labov 1990), but also in studies transecting time and culture (Blaxter 2015; Gal 1978; Haeri 1994; Knooihuizen 2016; Labov 2015; Trudgill 1983). It has been argued that gender is universally experienced as an important category for

https://doi.org/10.1515/9783110435351-021

organizing social interaction (Holmes 2007; Pavlenko 2001). We will suggest that the widespread absence of gender effects in language contact is predicted by Ochs' (1992) theory of indexicality, where she proposes that all social indexicalities are secondary to, or mediated by, direct indexicalities between a linguistic form and a stance, act, or activity. Because the mapping of stances, acts, and activities onto gender ideologies is culturally mediated, situations of language contact create conditions where (a) there may be more important social identities to inscribe linguistically, and (b) there may be competing ideologies about and ways of experiencing gender that need to be negotiated. Both of these considerations mean other social identities may be quantitatively more significant.

We do not want to reify the association between language and culture (aside from anything else, the notion of 'culture' is hardly without problems), and we do not claim that contact between languages necessarily entails contact between two cultures (obviously, there are some cultures where bi/multilingualism is a core cultural value; virtually all of Melanesia, South Asia, and Sub-Saharan African come to mind). However, there are comparatively few studies of gender and language in stable situations of language contact/multilingualism such as these. The paucity of such studies raises the possibility that the reason we have little evidence of gender as a significant constraint in language contact is because researchers have not looked for such evidence (absence of evidence being rather different than evidence of absence). Nevertheless, we will review some studies that have considered variation in stable situations of language contact, and see that they generally fail to show significant intergroup differences between female and male speakers.

We will suggest that there are structural reasons why gender seldom emerges as a significant basis for stratifying groups within a speech community in language/cultural contact. We propose that the principal sociolinguistic work undertaken in these contexts is: first, to establish an individual's personal stance(s) towards the indexicalities that different languages have in situations of language contact; and second, to establish the relationship between the speakers of the different languages in contact with one another. In situations where cultural or ethnic differences may be highly salient, we suggest that there are higher priorities than marking gender differences. We will suggest that, in high contact situations, complex renegotiations of social meanings and individual identities 'wash out' gender stratification. Inscribing gender differences as a significant predictor of language variation may be a luxury that speakers can focus on later. In short, gender may well be a universal category, but that doesn't entail that it is the primary basis for social organization. Hence, it also does not entail that gender will be the primary axis for sociolinguistic differentiation.

2. Evidence of gender effects in language contact

Pavlenko (2001) reviews a number of studies or accounts of language contact where researchers have observed broad differences in: (a) whether a language is (or is perceived to be) primarily used by women or men, and (b) the degree to which men and women engage in bi/multilingualism. If a language is (newly) introduced to a community through language contact, the choice to embrace or reject that language can become a

gendered one. Beyogle's (2015) study of French and English borrowings in Dagara, on the border of Burkina Faso and Ghana, found that French borrowings were more typical among women, and English ones among men.

Variation in preferred language of use is often associated with power, economic opportunity, or individual agency. The comparatively well-studied relationship between Arabic and French in Tunisia illustrates this. Sayahi (2011) finds no difference in frequency of French borrowings into women's and men's Arabic, but since the study included only highly educated women, it is impossible to tease apart whether education or gender affects the propensity to use French borrowings. Other work (Trabelsi 1991; Walters 2011) demonstrates that the use of French in Tunisia is gendered in complex ways, intersecting not only with other demographic labels (use of French and Standard French pronunciations are particularly associated with young, educated, urban women), but also with historically contingent facts about the practice of religion, attitudes towards marriage, and the development of a post-colonial national identity. French signifies education, international mobility, and/or cultural openness, and both women and men draw on that capital, depending on the details of their personal biography or their personal stance towards societal beliefs about cultural authenticity or openness (Walters 2011). Hmong women living in the USA assert different social identities through their patterns of multilingualism and language shift (Stanford 2010). Besnier (2003, 2004) argues that use of English (introduced through travel and family ties in the diaspora) is gendered, but that this gendering must be understood in relation to other values: English is often used when the speakers want to distance themselves from traditional Tongan cultural values (this is often the motivation for Tongan *fakaleiti* − men who live as women in the community), or from what is prescribed in Tongan society for social action (which motivates the use of English by women in second-hand markets).

The link between language and 'value(s)' may sometimes be quite literal. Dubois and Horvath's (2000) work on Cajun English shows that the use of traditional French features in Cajun English in Louisiana correlates with the relative economic and social value of French and English. Younger men working outside the local community use more French features in their English than do older working men or younger working women: young men who make their living in tourism and the music business apparently find the use of French has value in constructing their working personas.

Where bilingualism is valued for employment, and where work patterns are gendered, it is not surprising to find that language use is also gendered (Holmes 1993; Menard-Warwick 2006; Piller 2016). Again, though, the gendering of language choice is derived from social opportunities that have material costs and benefits. Aikio (1992) suggests that women may actively resist the shift towards a new language, if they perceive the power disparity between women and men to be greater in the outgroup than in their ingroup. That is, their avoidance of the language serves as a symbolic means of distancing themselves from the cultural values that they associate with speakers of that language. Winter and Pauwels (2000) outline other examples of the sometimes gendered nature of community language maintenance (see also James 1996).

Virtually all of these studies demonstrate the importance of nuanced, ethnographically-informed analysis. They show that individuals adopt positions in their everyday interactions that assert how they relate (or want to be perceived to relate) to dominant ideologies about the social meaning of different languages. Moreover, by adopting such stances towards those ideologies, they help to recreate or contest those ideologies.

Numerous case studies of language acquisition and language learning point to language as an important sociolinguistic resource. Through their linguistic behavior, speakers index their stance towards the values and norms that they associate with different languages. MacPherson's (2005: 585) account of five refugee learners of English describes "five alternate patterns of identity and language negotiations: rejection, assimilation, marginality, bicultural accommodation, and intercultural creativity". Contact with a new culture and language required that the nuns (re)negotiate and (re)frame their identities as women, as Buddhists, and as Tibetans. The nuns experienced the interrelation of culture, language, and gender in highly personal ways, but these different experiences of their macro-identities (woman, Buddhist, Tibetan) might never have become evident, had they not been forced into a language contact situation.

This is reminiscent of Menard-Warwick's (2004) study of two Latin-American migrants to the US. Despite their superficially similar backgrounds and even their similar overt ideologies about gender, their experiences led to very different language learning behaviors. Shi (2006) also demonstrates how gender can have a profound effect on a speaker's acquisition or use of a new language.

Walters (2011) shows how 'French-ness' (and the social values associated with French) influence the gendered use of: (a) borrowings in the lexicon, (b) the stereotyped pronunciation of [r] (Tunisian) versus [ʁ] (French), and (c) patterns of French/Arabic code-switching. Just as personal stances scale up to create social meaning for different languages, Walters (2011) shows that the locus of variation also scales up from the particular (phonemes) to the general (language choice).

Alongside the numerous qualitative studies showing that the use of different languages (or features associated with a language) may express an individual's stance or subject position in everyday discourse, we find an extensive (largely) qualitative body of research on language policy and planning, which also demonstrates that languages index various power and status differences between cultural or social groups. Some of this research notes that people's experience of power and status is mediated by gender; that is, women and men have differential access to power in different social spheres.

3. Weak evidence for gender stratification in communities

In quantitative studies of language contact, if gender stratification is reported, the effects of gender are unpredictable, and the results in many cases seem incompatible with the principles of gender and language variation/change derived from research on monolingual speech communities (Labov 2001, 2015).

Liu (2012) reports that women are leading men in a shift to the (stigmatized) pronunciation of the pronoun *wo* 'I' as [u:] in Shanghai. But other social factors are important in predicting the use of [u:] − these include the changing social meaning of this variant over time (from marking 'outsiders' to marking cosmopolitanism, see Zhang 2005), and speakers' networks/identification with recent migrants to Shanghai.

Dorian (2006) found that women were more likely than men to retain a conservative form of the vocative in Sutherland Gaelic, even while the community as a whole was shifting to English. There are also cases where gender shows no significant effect at all (Nagy 2011; Rau 2000). Gender effects seem to be much less likely to be marked quanti-

tatively than social class (Lynch 2009; Rau 2000), education (possibly a proxy for social class, Sayahi 2011), ethnicity (Winter and Pauwels 2000), and social networks (more on networks shortly).

Perhaps most tellingly, Barnes' (2009) investigation into the use of Veneto (a variety of Italian), isolated over time in a village in Mexico, fails to support previous claims that there are gender differences in the use of Veneto. The qualitative or subjective perceptions of gender difference do not show up in Barnes' quantitative analysis of patterns of language use. This is not to say that qualitative generalizations are wrong; it simply speaks to the complementary nature of qualitative and quantitative methods. This raises another issue, going well beyond the scope of this chapter, namely whether variation observed in situations of language contact is contact-induced or internally motivated change. Poplack (1996) and Poplack, Zentz, and Dion (2012) argue that many putative cases of contact-induced change are miscategorized instances of internal change.

We have already noted that the gender patterns observed in the Cajun community are meaningful in the context of speakers' employment and social networks. Dorian (2006), too, explained the use of traditional Gaelic vocatives in younger women's speech by their close network ties with older, Gaelic-dominant women in the community. In many variationist studies of language contact, it seems that social network membership has a far stronger influence on language use (or specific variables) than gender *per se*. Shin and Otheguy (2013) examined the use of overt subject pronouns in a number of different groups of Spanish-speaking migrants to the US. They find that among the most innovative speakers there is a gender effect: women use more overt pronouns in their Spanish than men do. Yet gender is a much weaker predictor than exposure to or contact with English. They suggest the gender effect is an artefact of female migrants having social networks that bring them into greater contact with US-born speakers of Spanish and speakers of English (Nichols 1983).

Lynch's (2009) study of /s/ lenition and deletion in Miami Spanish found that young, upper-middle class women produce more zero variants, and young, upper-middle class men produce more overt realizations of /s/ (which is very different from the way this variable patterns in varieties of monolingual Spanish). In addition, younger Miami speakers of Spanish use more /s/ overall than their grandparents do. Hence, Lynch suggests that the use of /s/ may be a way for younger speakers to overtly distance themselves from social stereotypes about their grandparents' generation of migrants.

These examples reflect situations of dynamic change in the linguistic resources of a speech community, e.g. through migration or other cultural upheaval. But even in stable contexts of language contact, where principles of language variation and change might be expected to apply (such as Labov's principle that women lead changes occurring below the level of conscious awareness), researchers often fail to find any evidence of intergroup differences according to gender. Studies of language variation in Bislama and some indigenous languages of Vanuatu (Meyerhoff 2000, 2002, 2015) fail to find systematic gender differences. Labov's (1993) formulation of the sociolinguistic monitor would predict that (morpho)syntactic variables, such as make up the bulk of Meyerhoff's work, will not show social stratification. However, there is no gender effect for lexical alternations either, despite the high social awareness attached to them (Meyerhoff 2016). Bequia (St. Vincent and the Grenadines) can be treated as the stable outcome of historical language contact, and also as the site of stable, ongoing contact between the local creole, other Caribbean creoles, and Standard (Caribbean) English. As in Vanuatu, extensive

study of Bequia creole consistently finds that sociolinguistic variables − whether morphosyntactic (Meyerhoff and Walker 2013), phonological, or lexico-syntactic (Walker and Sidnell 2011) − show some social stratification (notably speaker's village of origin), but never gender stratification (Meyerhoff 2008).

4. Speculations on the gap

In order to offer some motivation for this gap, we adopt Ochs' (1992) architecture for analyzing language and social meaning (principally her distinction between *direct* and *indirect* indexicality). Silverstein's (2003) orders of indexicality is related, however, his approach does not preclude gender being a first order indexicality. Our review of the language contact case studies suggests that Ochs' (1992) model better describes the data. Ochs claims that gender is always an indirect index derived from personal stances or social acts. In brief, Ochs (1992) distinguishes direct markers or indices of gender, from the vast majority of (probabilistically distributed) linguistic features that people like to think of as being markers of gender. Linguistic forms that are a direct index of gender include pronouns or kinship terms (that is, words where a feature of [+female/+male] is part of the truth-conditional semantics of the word). Linguistic forms such as tag questions, hedges, and sentence-final particles derive their gendered meaning indirectly. Ochs (1992) argues that what such forms directly index is not gender, but rather personal stances (e.g. in Tunisia, being open or traditional) or social acts and activities (such as hawking second-hand goods in Nuku'alofa markets). To the extent that the community believes that a particular stance or activity is more properly associated with women versus men, the linguistic forms which directly index that stance/activity indirectly acquire a gendered meaning. This means, therefore, that gender is always secondary or derived.

Holmes (2007: 56−57) argues that, the cultural turn in sociolinguistics notwithstanding, "most of the world continues to treat 'women' and 'men', 'female' and 'male' as fundamental social categories", and it is possible that in situations of language contact, gender is highly accessible and ripe for co-optation when a language learner is faced with the complexity of inter- and intra-speaker variation. Meyerhoff and Schleef (2012) compare the constraints on variation between [ɪn] and [ɪŋ] in British English and find that migrant, non-native speakers match the locally-born teens in many respects, but in London, they also innovate a gender constraint which is not modelled by their locally-born peers. Meyerhoff and Schleef (2012) suggest that identifying and matching native speaker constraints on variation is a highly complex task, and this complexity allows a fundamental social category like gender to emerge as a statistically significant constraint among non-native speakers.

Notably, however, by their analysis, gender is not the only fundamental category. Social and friendship networks − and not gender − show up as the only statistically significant social constraint for migrants to Scotland. Some social structure seems to be imposed, but what is most salient is contextually determined.

It seems reasonable to believe that, in situations of language contact − especially those caused by migration, which constitute the bulk of the data reviewed here − if speakers are searching for ways to imbue their speech with social meaning, social catego-

ries other than gender may be uppermost. Differences in ethnicity, who you socialize with, and your social class within society all probably have much more immediate impact on your everyday interactions, on how you imagine your life chances, and on the aspirations you have for your children. Translating this to Ochs' (1992) paradigm, the indirect indexicalities that accrete from stances and activities will be more likely to be interpreted (first) in terms of these social categories rather than gender. We would not expect, therefore, to find a significant gender effect on a variable, unless there is also evidence that the variant favored by one gender is also favored in exactly those social activities or interpersonal stances that are socially constructed as gendered in that community. The fact that gender emerges as a weak constraint in Shin and Otheguy (2013) perhaps speaks to the relative time depth of Hispanic migration to the US; the fact that either gender or friendship networks emerge as significant indirect indexes among high school teenagers (Meyerhoff and Schleef 2012) perhaps speaks to the immediate and material relevance of these social categories in adolescence.

Many people view the world through a gendered lens, and there is no doubt that gender continues to have material consequences in how people live their lives. In sociolinguistics, it has been shown to be a recurring constraint on individuals' life chances and social action. It has also been implicated as playing a central role in the transmission of variation and its transformation into lasting change. We have seen that, in situations of language contact, gender roles and gender ideologies interact with and shape individuals' behavior and provide a frame for others' interpretations of a speaker's linguistic behavior. In some cases, we have seen that they play a role in shaping intergroup behavior.

However, we have also noticed that in language contact, gender seldom plays a starring role in the stratification of variation and change. In fact, in all but one of the case studies we examined, researchers pointed out that gender needed to be interpreted in light of other macro-social categories, such as social networks, class, or ethnicity. Some of the reasons underpinning this may be methodological: the Labovian generalizations are by and large (though see Labov 2015) based on studies of Western, monolingual speech communities where there is a standardized, normative variety. There may also be issues with structuralist assumptions underlying the social categories investigated in such communities (class, ethnicity, gender).

We suggested an alternative perspective: namely, that the absence of gender as a significant constraint on intra-individual variation in situations of contact is not entirely surprising if we adopt Ochs' (1992) perspective on the association between linguistic forms and social meaning. Forms (or linguistic codes) directly index stances or subject positions; hence, gender can always be expected to emerge in qualitative studies of individuals' experience. Broader social categories emerge only as secondary associations that are dependent on stances.

In many of the contexts where languages come into contact, social identities other than gender are of greater practical, everyday importance. We might expect to find more evidence of gender stratification as situations of language contact stabilize and settle down, though the limited evidence suggests that, even there, gender is seldom (or never) a significant basis for social stratification of variation. This would suggest that, where language contact is transforming into language shift, we might be much more likely to find gender emerging as a significant basis for constraining variation within and between languages. Clearly, there is a great need for more study of the ontology of gender in language contact communities.

5. References

Aikio, Marjut
1992 Are women innovators in the shift to a second language? A case study of Reindeer Sami women and men. *International Journal of the Sociology of Language* 94. 43–61.

Barnes, Hillary
2009 *A sociolinguistic study of sustained Veneto-Spanish bilingualism in Chipilo, Mexico.* University Park, PA: The Pennsylvania State University dissertation.

Besnier, Niko
2003 Crossing genders, mixing languages: The linguistic construction of transgenderism in Tonga. In Janet Holmes & Miriam Meyerhoff (eds.), *The handbook of language and gender*, 279–301. Oxford: Blackwell.

Besnier, Niko
2004 Consumption and cosmopolitanism: Practicing modernity at the second-hand marketplace in Nuku'alofa, Tonga. *Anthropological Quarterly* 77(1). 7–45.

Beyogle, Richard
2015 *Language contact in two border communities in Burkina Faso and Ghana: Lexical borrowings from French, English and African Languages*. Urbana, IL: University of Illinois at Urbana-Champaign dissertation.

Blaxter, Tam
2015 Gender and language change in Old Norse sentential negatives. *Language Variation and Change* 27(3). 349–375.

Cameron, Deborah
2014 Gender and language ideologies. In Susan Ehrlich, Miriam Meyerhoff & Janet Holmes (eds.), *The handbook of language, gender and sexuality*, 281–296. Oxford & New York: Wiley-Blackwell.

Dorian, Nancy C.
2006 Negative borrowing in an indigenous language shift to the dominant national language. *International Journal of Bilingual Education and Bilingualism* 9(5). 557–577.

Dubois, Sylvie & Barbara Horvath
2000 When the music changes, you change too: Gender and language change in Cajun English. *Language Variation and Change* 11(3). 287–313.

Eckert, Penelope
1989 The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1(3). 245–67.

Gal, Susan
1978 Peasant men can't get wives: Language change and sex roles in a bilingual community. *Language in Society* 7(1). 1–16.

Haeri, Nilofar
1994 A linguistic innovation of women in Cairo. *Language Variation and Change* 6(1). 87–112.

Hellinger, Marlis & Hadumod Bußmann
2001 *Gender across languages*. Amsterdam & Philadelphia: John Benjamins.

Holmes, Janet
1993 Immigrant women and language maintenance in Australia and New Zealand. *International Journal of Applied Linguistics* 3. 159–179.

Holmes, Janet
2007 Social constructionism, postmodernism and feminist sociolinguistics. *Gender and Language* 1(1). 173–183.

James, Deborah
1996 Women, men and prestige speech forms: A critical review. In Victoria Bergvall, Janet Bing & Alice Freed (eds.), *Rethinking language and gender research: Theory and practice*, 98–125. London & New York: Longman.

Knooihuizen, Remco
2016 Convergence in generic pronouns: Language contact and Faroese *mann*. *Acta Linguistica Hafniensia* 47(2). 220–243.
Labov, William
1990 The intersection of sex and social class in the course of linguistic change. *Language variation and change* 2(2). 205–254.
Labov, William
1993 The unobservability of structure and its linguistic consequences. Paper presented at the conference New Ways of Analyzing Variation 22, University of Ottawa, 14–17 October.
Labov, William
1994 *Principles of linguistic change, vol. 1: Internal factors.* Oxford: Blackwell.
Labov, William
2001 *Principles of linguistic change, vol. 2: Social factors.* Oxford: Wiley-Blackwell.
Labov, William
2010 *Principles of linguistic change, vol. 3: Cognitive factors.* Oxford: Wiley-Blackwell.
Labov, William
2015 The discovery of the unexpected. *Asia-Pacific Language Variation* 1(1). 7–22.
Liu, Guo-Qiang
2012 Social identity and sound change. *Australian Review of Applied Linguistics* 35(2). 203–214.
Lynch, Andrew
2009 A sociolinguistic analysis of final /s/ in Miami Cuban Spanish. *Language Sciences* 31(6). 766–790.
MacPherson, Seonaigh
2005 Negotiating language contact and identity change in developing Tibetan-English bilingualism. *TESOL Quarterly* 39(4). 585–607.
Menard-Warwick, Julia
2004 "I always had the desire to progress a little": Gendered narratives of immigrant language learners. *Journal of Language, Identity, and Education* 3(4). 295–311.
Menard-Warwick, Julia
2006 'The thing about work': Gendered narratives of a transnational, trilingual Mexicano. *International Journal of Bilingual Education and Bilingualism* 9(3). 359–373.
Meyerhoff, Miriam
2000 The emergence of creole subject-verb agreement and the licensing of null subjects. *Language Variation and Change* 12(2). 203–230.
Meyerhoff, Miriam
2002 Formal and cultural constraints on optional objects in Bislama. *Language Variation and Change* 14(3). 323–346.
Meyerhoff, Miriam
2008 Men argue, but the women doz trace: Gender and language variation in Bequia (St Vincent and the Grenadines). *Sargasso 2008–09* 1. 115–132.
Meyerhoff, Miriam
2015 Turning variation on its head: Analysing subject prefixes in Nkep (Vanuatu) for language documentation. *Asia-Pacific Language Variation* 1(1). 79–109.
Meyerhoff, Miriam
2016 Borrowing from Bislama into Nkep (East Santo, Vanuatu): Quantitative and qualitative perspectives. *Languages and Linguistics in Melanesia.* 34(1). 77–94. http://www.langlxmelanesia.com/issues (accessed 6 September 2017).
Meyerhoff, Miriam & Erik Schleef
2012 Variation, contact and social indexicality in the acquisition of (ing) by teenage migrants. *Journal of Sociolinguistics* 16(3). 398–416.

Meyerhoff, Miriam & James A. Walker
2013 *Bequia Talk: St. Vincent and the Grenadines*. Westminster: Battlebridge Publishing.
Nagy, Naomi
2011 Lexical change and language contact: Faetar in Italy and Canada. *Journal of Sociolinguistics* 15(3). 366–382.
Nagy, Naomi & Miriam Meyerhoff
2008 Introduction: Social lives in language. In Miriam Meyerhoff & Naomi Nagy (eds.), *Social lives in language: Sociolinguistics and multilingual speech communities*, 1–16. Amsterdam & Philadelphia: John Benjamins.
Nichols, Patricia C.
1983 Linguistic options and choices for Black women in the rural South. In Barrie Thorne, Cheris Kramarae & Nancy Henley (eds.), *Language, Gender and Society*, 54–68. Cambridge, MA: Newbury House.
Ochs, Elinor
1992 Indexing gender. In Alessandro Duranti & Charles Goodwin (eds.), *Rethinking context: Language as an interactive phenomenon*, 335–358. Cambridge: Cambridge University Press.
Pavlenko, Aneta
2001 Bilingualism, gender, and ideology. *International Journal of Bilingualism* 5(2). 117–151.
Piller, Ingrid
2016 *Linguistic diversity and social justice: An introduction to applied sociolinguistics.* Oxford & New York: Oxford University Press.
Poplack, Shana
1996 The sociolinguistic dynamics of apparent convergence. In Gregory R. Guy, Crawford Feagin, John Baugh & Deborah Schiffrin (eds.), *Towards a social science of language: Papers in honor of William Labov*, 285–309. Amsterdam & Philadelphia: John Benjamins.
Poplack, Shana, Lauren Zentz & Natalie Dion
2012 Phrase-final prepositions in Quebec French: An empirical study of contact, code-switching and resistance to convergence. *Bilingualism: Language and cognition* 15(2). 203–225.
Rau, Der-Hwa. V.
2000 Phonological variation and sound change in Atayal. *Oceanic Linguistics* 39(1). 144–156.
Sayahi, Lotfi
2011 Code-switching and language change in Tunisia. *International Journal of the Sociology of Language* 211. 113–133.
Shi, Xingsong
2006 Gender, identity and intercultural transformation in second language socialisation. *Language and Intercultural Communication* 6(1). 2–17.
Shin, Naomi Lapidus & Ricardo Otheguy
2013 Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York. *Language in Society* 42(4). 429–452.
Silverstein, Michael
2003 Indexical order and the dialectics of sociolinguistic life. *Language and Communication* 23(3–4). 193–229.
Stanford, James
2010 The role of marriage in linguistic contact and variation: Two Hmong dialects in Texas. *Journal of Sociolinguistics* 14(1). 89–115.

Trabelsi, Chedia
1991 De quelques aspects du langage des femmes de Tunis. *International Journal of the Sociology of Language* 87. 87–98.
Trudgill, Peter
1983 Social identity and linguistic sex differentiation. In Peter Trudgill, *On Dialect*, 161–168. Oxford: Blackwell.
Walker, James A. & Jack Sidnell
2011 Inherent variability and coexistent systems: Negation on Bequia. In Lars Hinrichs & Joseph Farquharson (eds.), *Variation in the Caribbean*, 39–55. Amsterdam & Philadelphia: John Benjamins.
Walters, Keith
2011 Gendering French in Tunisia: Language ideologies and nationalism. *International Journal of the Sociology of Language* 211. 83–111.
Winter, Joanne & Anne Pauwels
2000 Gender and language contact research in the Australian context. *Journal of Multilingual and Multicultural Development* 21(6). 508–522.
Zhang, Qing
2005 A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society* 34(3). 431–466.

Miriam Meyerhoff, Wellington (New Zealand)
Alexandra Birchfield, Wellington (New Zealand)

III. Societal aspects of language contact

22. Multilingualism

1. Introduction
2. Defining multilingualism
3. Distribution and incidence of multilingualism
4. Changing face of multilingualism and superdiversities old and new
5. Conclusions and directions for future research
6. References

1. Introduction

Multilingualism is a complex, multifaceted phenomenon manifesting itself in various ways to differing degrees within countries, communities, institutions, families, and individual brains. Researchers offer many definitions of multilingualism, sometimes distinguishing it from terms like 'bilingualism' and 'plurilingualism'. Some use the term 'bilingual' for users of two languages/varieties and 'multilingual' for users of three or more. Some scholars believe that bilingualism and multilingualism differ in terms of complexity and should remain distinct fields, each employing their own research methods and populations (Aronin and Jessner 2015: 60). Nevertheless, the term 'multilingualism' is increasingly used as a generic, inclusive term referring to use of two or more languages/varieties (Cenoz 2013: 7). Some also distinguish between 'multilingualism' as the presence of many languages/varieties in society (i.e. societal multilingualism) and 'plurilingualism' as presence of many languages/varieties in individuals (i.e. individual multilingualism). Not everyone uses key terms in the same way, which makes comparing studies difficult.

This chapter regards multilingualism as a relative concept involving at least two or more languages/varieties; it makes no distinction between bilingualism and multilingualism as separate concepts. Multilingualism is an inherently relative notion, because it is impossible to define non-arbitrarily when a person/society is bi- or multilingual. Classic definitions of bilingualism focused on opposite ends of what is better regarded as a continuum. Maximalist ones like Bloomfield's (1933: 56) required "native-like control of two languages", while minimalist ones like Diebold's (1961) 'incipient bilingualism' allowed those with minimal knowledge in at least one of the four basic components of language competence (i.e. speaking, listening, reading, writing) to be considered bilingual. The former is clearly too restrictive and the latter too broad. Because it is not always possible to maintain strict boundaries between individual and societal multilingualism, this chapter does not distinguish between plurilingualism and multilingualism. The presence of more than one language in the same geographical area does not entail that all inhabitants know both languages. Officially monolingual territories or countries, like France, may have sizable multilingual populations, while monolinguals often live in multilingual societies and countries, like Switzerland. Nevertheless, since part III of Volume I in this handbook concerns 'societal aspects of language contact', this chapter adopts a predominantly macrosociolinguistic perspective, concentrating primarily on societal rather than individual multilingualism.

https://doi.org/10.1515/9783110435351-022

After defining some key terms and concepts in Section 2 of this chapter, Section 3 offers an overview of types of multilingual communities of varying sizes, ranging from small scale societies to large ones arising from more complex forms of sociopolitical organization (e.g. multilingual empires, nations, and cities), as well as those arising from interactions in so-called 'communities of practice', both in the real world and online. Section 4 considers some contemporary global forces changing the character of multilingualism as migration intensifies linguistic heterogeneity in many places and English increasingly becomes a component of multilingual repertoires. The chapter concludes with a short summary of key points and a look at future areas of research and challenges.

2. Defining multilingualism

No precise or comprehensive statistics exist concerning the distribution and incidence of multilingualism, either historically or currently, but it is by no means a new phenomenon in human history. Patterns of multilingualism and types of multilingual communities arise from various kinds of language contact and factors, e.g. political union, conquest, marriage, trade, migration, schooling, and globalization. In today's globalized world, being multilingual is a normal state of being for as many as three quarters of the world's population, who use at least two languages in their everyday lives, as well as perhaps as many as half, who use at least three (Baker and Prys Jones 1998: vii).

Despite the normality of multilingualism and the centrality of language contact to the evolution of language, research on multilingualism has occupied a relatively marginal role within linguistics until recent decades. No doubt influenced by the focus of linguistic theory on Chomsky's (1965: 3) abstraction of the "ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly", monolingualism has been assumed to be the natural human condition, the unmarked case to which multilingualism is compared. This predominantly monolingual orientation can be seen in the fact that numerous books offer overviews and surveys of multilingualism, but "it would certainly be odd to encounter a book with the title, *Monolingualism*" (Romaine 1995: 1). Against the backdrop of a predominantly monolingual ideology, ideas about multilingualism have been adversely influenced by terms like 'ideal', 'full', and 'balanced', which assume other forms of bilingualism exist that are not ideal, full, or balanced. Early studies tended to elevate 'ideal bilinguals' to some special theoretical status, as seen in Bloomfield's (1933: 56) restriction of bilinguals to persons with "native-like control of two languages". Even today, multilinguals are still largely assessed by reference to monolingual norms: the vast majority of measures developed for bilinguals are the same ones used for monolinguals.

The hegemony of monolingualism has also meant that multilingualism has been understood as multiple monolingualisms in distinct languages, mastered separately in standard form and kept pure of outside influence. Measured by this yardstick, mixed ways of speaking common to many multilingual communities have been regarded as somehow deviant − the result of imperfect competence, if not downright ignorance and laziness − rather than as legitimate modes of communication in their own right. Accepting Weinreich's ([1953] 1968: 73) characterization of the ideal bilingual who "switches from one language to another according to appropriate changes in the speech situation (interlocu-

tors, topics, etc.), but not in an unchanged speech situation" would entail concluding that most multilinguals are not ideal and have less than native-like control of the languages they speak. In everyday life, degree and extent of multilingualism depend on Fishman's (1965) basic question of who speaks what to whom and when. Speakers are not always able to choose the language(s) they prefer. Multilinguals operate along a continuum, where different modes of speaking are both constrained and induced by context. At one end, they may be in totally monolingual mode if interacting with monolinguals of one or other languages they know, while at the other, they may be in multilingual mode when communicating with others who share the same two (or more) languages and with whom they normally mix, switch, and borrow (Grosjean 1997: 228).

Defining a complex phenomenon like multilingualism is also rendered problematic by the assumption that languages exist as discrete, separate, and countable entities. Indeed, a central problem of linguistics is "how to reconcile a convenient and necessary fiction with a great mass of inconvenient facts" (Strevens 1982: 23), manifested in the actual diversity of language use. In the heterogeneous communities and contexts where multilingualism prevails, individuals are rarely equally fluent in the languages they know. This makes it impossible to define where multilingualism begins, or to specify non-arbitrarily the minimal proficiency for bilingualism. Individual repertoires may include multiple languages with differing acquisition histories, and varying proficiency levels and patterns of use across a lifespan. Some multilinguals grow up with two or more languages simultaneously from birth, while others acquire additional languages sequentially in different contexts later, e.g. school, work, etc.

3. Distribution and incidence of multilingualism

Multilingualism is possibly as old as the human species itself; the primal human condition sustained linguistic diversity for most of human history, during which people lived in small communities, largely as hunter-gatherers or small-scale agriculturalists. This lifestyle tended to prevent the emergence of dominant languages and cultures due to the relative social homogeneity of small groups, whose size was dictated by the number of people supportable by local resources. Communities were always in contact with others and connected to them economically and/or socially through exchange of goods, knowledge, and marriage partners. After the Neolithic period, which was ushered in by the invention and spread of agriculture about 12,000 years ago, created conditions for the rise of large groups whose spread was stopped only by natural barriers and ecologies unsuitable for farming, larger and more complex types of political communities like cities, empires, and states evolved. Over the last 300 years, the prevailing polity has been the sovereign state, which emerged from previously existing empires, to become the bedrock of the current political world order.

3.1. Small-scale multilingualism

In some parts of the world, hotspots of linguistic diversity survive where multilingual households are still the norm. In Morehead District of Papua New Guinea's Western

Province, bordering the southern coast and the border between Indonesian New Guinea, exogamous marriages bring wives and their languages into villages like Bimadbn. Villagers claim Nen, with circa 250 speakers, as their village language, but they are also fluent in Nmbo (a related language), Idi (an unrelated language), and the regional lingua franca, Motu. Some also know more recently introduced languages, like Tok Pisin, an English-based pidgin and widespread lingua franca, and English. Daughters marry into their mother's village, ensuring a continuing cycle of multilingualism. By contrast, on the Indonesian side of the border, economic development and environmental change are proceeding rapidly; speakers of traditional languages are now significantly outnumbered by transmigrants from elsewhere in Indonesia. Language shift from local languages to Indonesian as a dominant language is already reaching a critically advanced state among younger people (Evans 2012: 119–121).

In Misión La Paz, Argentina, three indigenous languages (Chorote, Nivaclé, and Wichí) coexist with Spanish, introduced through colonialism. Linguistic exogamy results in passive rather than reciprocal multilingualism. Although individuals identify with one language, which they speak to everyone else, they claim to only understand, but not speak, other languages spoken to them. Thus, husbands and wives typically speak different languages to one another, and children in the same family very often speak and identify with languages which are different from one another (Campbell and Grondona 2010). Warruwi, a small community of circa 400 located on South Goulburn Island in northwest Arnhem Land, Australia, also maintains high levels of individual and societal multilingualism. Adults commonly speak two to four languages, but also understand many more; most children can already speak two indigenous languages when they start school. Each person speaks to the other in their preferred language, while understanding the other's language fairly well (Singer and Harris 2016).

3.2. Multilingual empires

Before massive differences existed between the expansionary potential of different peoples of the type that might cause a single, dominant language group to spread over a large area, egalitarian multilingualism could flourish in small communities with traditional subsistence economies. The relationship between spread of languages and empires was established centuries before Antonio de Nebrija, Bishop of Avila ([1492] 1989: 109), observed in the introduction to his Spanish grammar: "*Siempre la lengua fue compañera del imperio, y i de tal manera lo siguió que juntamente comenzaron, crecieron, i florecieron, y después junta fue la caída de ambos*" ['Language has always been the companion of empire and in such a way it followed that together they began, grew, and flourished, and afterwards together they fell']. As the site of the oldest farming cultures, the ancient Near East well illustrates some of the earliest associations between the fates of languages and empires when the oldest towns and cities consolidated their power into empires. After conquering Sumeria's city-states, Sargon I created the first multinational empire, centered in the city of Akkad, which he ruled from 2334 to 2279 BCE. At the pinnacle of the Assyrian Empire's power in the 8th century BCE, Aramaic replaced Akkadian, turning a previously largely monolingual Assyrian society into a multilingual one. A ruling class bilingual in Akkadian and Aramaic governed a multiethnic state

comprising many peoples and tribes of different origins, with numerous Assyrian citizens identifying as Egyptians, Israelites, Arabs, Anatolians, and Iranians. From the 7th until the 4th century, Aramaic served a succession of great empires (the Assyrian, the Babylonian, and the Persian), until Alexander the Great's conquest ended the Persian Empire and made Greek the imperial lingua franca, until it in turn was replaced by Latin in the western Roman Empire (Parpola 2004). Today, however, Aramaic, once a tool for conquest, commerce, and government, for consolidating political power, and for spreading religions and cultures across a vast empire stretching from Egypt to India and China, is spoken by only about half a million Assyrians in Iraq, northeastern Syria, southeastern Turkey, northwest Iran, and diaspora communities in Armenia, Georgia, Azerbaijan, southern Russia, and elsewhere.

The Roman Empire was also linguistically diverse. Despite extending Latin as the language of administration and citizenship, as well as of military and legal rule across the Mediterranean, North Africa, and Britain, it did not replace Greek; local languages were also used alongside Latin and Greek, and some survived the end of Roman rule (Adams 2003). Just as Roman educated elites were bilingual in Greek, ruling monarchs in subsequent empires were multilingual, too. The constitution for the Holy Roman Empire promulgated by Charles IV in the Golden Bull of 1356 advised emperors and princes to be taught by the age of seven to speak the Latin, Italian, and Slavic tongues in addition to their own High-German mother tongues so that they could communicate with their subjects. Charles V, Holy Roman Emperor from 1519 to 1556 (Charles I of Spain 1516–1556), is reported to have said that he spoke Spanish to God, Italian to women, French to men, German to his horse, and English to birds (Burke 2004: 28). By that time, his dominion encompassed numerous overseas Spanish colonies and territories in the New World, where Spanish was imposed by military force on already existing multinational and multilingual empires. Nahuatl, the language of the dominant tribe ruling over the Aztec empire covering much of what is now Mexico and Central America, functioned as official language and lingua franca for as many as 80 language groups (Ostler 2005). When Spanish colonization of the Americas began, ushering in a new phase of global language expansion by sea, Spain itself had just emerged from some 700 years of Moorish occupation as a politically united nation for the first time, bringing together different groups under one language (Castilian) and one religion (Catholicism).

3.3. Multilingual states

The first modern states arose and consolidated themselves from the dissolution of large empires, like the multilingual, multinational Holy Roman Empire, dramatically changing the nature of territorial organization across the world and transforming patterns of societal and individual multilingualism. As the idea of 'one nation-one language' developed during the late 18th and early 19th centuries in the context of European nation-building, a common national language came to be seen as inseparably bound to national culture, a sign of loyalty to the nation itself, and essential for modernization and economic progress. With approximately 7,000 languages today (Simons and Fennig 2017), but only about 200 nations, there are about 35 times as many languages as countries. Bilingualism or multilingualism is present in practically every nation, whether officially

recognized or not (Romaine 1995). Despite the fact that languages do not respect geopolitical borders and very few European states were ethnically or linguistically unified, the nation-state is arguably a (if not, the most) critical unit of analysis for studying societal multilingualism; policies pursued within national boundaries give some languages (and their speakers) the status of majority and others that of minority.

The linguistic heterogeneity of many countries reflects the arbitrariness of shifting political boundaries encapsulating distinct ethnic groups or nationalities with their own languages, cultures, and identities. Minority language communities may be split across one or more countries (e.g. Kurds in Turkey, Iraq, Iran, Armenia, and Syria, or Saami in Russia, Finland, Sweden, and Norway). Alternatively, members of the minority community in one state may form part of the majority community in a neighboring one (e.g. Finnish speakers in Sweden and Swedish speakers in Finland). Most of the territory of present-day Finland, for instance, was part of the Kingdom of Sweden from the 13th century to 1809, making Finnish a minority language both demographically and functionally. Until the late 19th century, the capital city of Helsinki was predominantly Swedish-speaking, and Swedish was the language of the elite, administration, and education. After Finnish independence in 1917, the fate of Finnish changed dramatically when it became the national language. Other languages, however, like signed languages used among deaf people, are minority languages in all contexts. Many minority languages have larger numbers of speakers than official state languages. Some newly independent multilingual countries in Africa and Asia have followed the same path taken in European nation-building by privileging languages of dominant ethnic groups. In Ethiopia, with more than eighty languages, more than a third of Ethiopians speak Oromo, but Amharic, spoken by only 29 % of the population as a first language, is the language of the dominant group and serves as the official language. In multiethnic-multilingual Malaysia, with 134 languages, the sole official and national language is Malay, which replaced the former colonial language, English.

Multilingualism poses challenges to monolingual ideologies associating monolingualism with nationhood. Only a quarter of all nations recognize more than one language (Edwards 2007: 44), and the majority of countries actually operate either *de facto* ('by fact') or *de jure* ('by law') as monolingual in recognizing only one language for use in education, administration, etc. In the United States, which has no official language policy, the predominance of English, spoken by just over 80 % of the largely monolingual population, obscures the multilingual reality of a territory that has been multilingual since time immemorial. With approximately 350 languages, it is the world's fifth most linguistically diverse country, and currently the fifth largest Spanish-speaking nation.

The European Union recognizes 24 national languages of its 28 member-states as official, but most member-states endorse official monolingualism at the national level; only six (Belgium, Finland, Ireland, Cyprus, Luxembourg, and Malta) are officially bi- or multilingual. The mother tongue of the majority of Europeans is one of the state languages of their country, but at least 40 to 50 million citizens of Europe speak languages other than the official language(s) of the state of which they are citizens. This does not include (im)migrant languages, whose speaker numbers are difficult to estimate. In most of them, however, there are minorities (both indigenous and non-indigenous), whose languages do not have the same rights as those granted to official languages.

In contrast to the egalitarian multilingualism found in small-scale, traditional societies, where language use is not primarily motivated by power relations or the prestige of

particular languages, within nation-states, the privileging of dominant languages typically results in hierarchical, asymmetrical forms of multilingualism like diglossia or polyglossia, a common form of societal bilingualism where two or more varieties/languages are functionally compartmentalized, each with its own domains of use. Typically, the so-called Low language/variety is used at home and in other personal domains of interaction between friends and family, while the High is used in public ones, like government, media, and education. When diglossia is stable, each language/variety has its own set of functions and spaces without threatening the other. Bilingualism without diglossia, however, tends to be transitional and results in language shift (Fishman 1991). In Europe, languages like English, Spanish, and French became symbols of the modernity of the newly-emergent nation-states, associated with higher social status, while other languages within their territories, like Scottish Gaelic, Basque, and Breton, respectively, became minoritized and stigmatized; their speakers often abandoned them in pursuit of social mobility.

3.4. Multilingual cities

In the 21st century, most of the world's population now lives for the first time in history in large metropolitan areas. As a modern, Asian multiethnic and multilingual city-state of over five million, Singapore offers a contemporary comparative perspective to some of the earliest city-states, like Akkad, as well as to western European nations and their increasingly diverse capitals, like London. Singapore has always been a multilingual and multiethnic society: Chinese currently comprise nearly 75 % of the population, Malays circa 13 %, Indians circa 9 %, and others circa 3 %. Centuries before Britain founded a trading port on the island in the early 19th century, Singapore was a meeting place for Arabs, Chinese, Malays, and traders from Indonesian islands and elsewhere in Southeast Asia, who brought various languages with them. After independence from Malaysia in 1965, the first language of the Singaporean Chinese community was predominantly Hokkien (followed by Teochew, Cantonese, Hainanese, and other Chinese languages), while the more linguistically homogeneous Malay community was largely Malay-speaking (85 %); the Indian community spoke mainly Tamil (59 %), Malayalam, and other languages. Contact between English and other languages gave rise to a new contact variety 'Singlish' (Singapore Colloquial English) which coexists alongside Standard Singapore English.

Unlike Malaysia, which gave preference to Malays, the dominant ethnic group, Singapore's newly independent government declared four official languages (Mandarin, English, Malay, and Tamil) to avoid privileging one group. The British colonial government used English, a practice the new government continued primarily for administrative purposes. The government focused language policy first on Mandarin for uniting the Chinese population through a "Speak Mandarin" campaign launched in 1979, and more recently on 'good' (i.e. Standard) English through a "Speak Good English" campaign begun in 2000. The intent of these policies was to make Mandarin the language of intergroup solidarity among Chinese, and the carrier of Asian rather than western values, while using English as a 'neutral' lingua franca for interethnic communication, but not as a language of identity. More recently, however, English is increasingly becoming the

mother tongue for more Singaporeans, as a younger generation educated in English-medium schools is increasingly more competent in English than their ethnic language(s). Most Singaporeans 45 years old and younger are generally at least functionally bilingual in English and one other official language (Bolton and Ng 2014).

London, too, was historically a multilingual and multicultural settlement from its very beginnings. Although the mix of cultures and languages has changed over time since its founding as a Roman city in 43 CE, it is by far the most ethnically and linguistically diverse part of the UK. As the largest city in Europe with a population of over 8 million, London is the number one destination for new migrants, and also home to some of Britain's longest-settled immigrant communities. Sizeable groups of European immigrants, merchants, traders, sailors, and craftsmen came to medieval and early modern London, and during the 17th, 18th, and 19th centuries as the British Empire expanded. African and Asian communities came as traders, soldiers, sailors or slaves and servants, often settling in distinct neighborhoods. After the Second World War, government and businesses started to recruit migrant workers, first from Europe, and then from many former, newly-independent colonies. Spitalfields in London's East End presents a snapshot of the city's ethnolinguistic diversity throughout the centuries: historically labelled 'Petty France' and 'Little Jerusalem', it is now called 'Banglatown' due to the large presence of Londoners of Bangladeshi origin.

In recent decades, London's population has become increasingly ethnically and linguistically diverse. London's first Muslim mayor, Sadiq Khan, whose parents moved from Pakistan to London in the 1960s, was elected in 2016 shortly before Britain voted to leave the European Union. Khan launched his "London is Open" campaign to reassure Londoners and visitors that the city would continue to have a diverse, multilingual population. Although London's first language overwhelmingly remains English (spoken as first language by 78 % of residents), and most Londoners are monolingual, Khan noted that over 300 languages were spoken in London. A promotional video celebrating the multilingual staff on London's transit networks featured a diverse sample of Londoners declaring in various languages that London is open, finishing with a London cabbie, speaking in English (Mehmedbegović et al. 2015).

3.5. Multilingual communities of practice

The concept of 'community of practice' provides another method for investigating community and societal multilingualism. Multilingual communities of practice include families, clubs, social networks, and more recently, online cybercommunities. The community's identity is defined by a shared domain of interest and a shared competence, distinguishing members from others (Wenger 1998). Not merely a club of friends or a network of connections between people, a community of practice interacts regularly and shares ways of doing things, ways of talking, beliefs, and values. An example specifically focused on language is Shaws Road (Bóthar Seoighe) Gaeltacht, a local Irish-speaking nucleus, in West Belfast, Northern Ireland. The community was created in 1969 when five families, who had acquired Irish and wanted to develop it as a family and community language, decided to build their houses together on a new housing development so that they could raise their children in an Irish-speaking environment, despite being sur-

rounded by a primarily monolingual English city where Irish had not been spoken for many years (Armstrong 2012).

Communities of practice do not, however, need physical space to interact. The digital revolution has created new, diverse forms of computer-mediated communication, e.g. email, instant messaging, discussion forums, blogs, and social networks like Facebook, YouTube, Twitter, etc. By connecting people across time and space, the internet provides opportunities for intercultural and multilingual contact in virtual reality, in real time or asynchronously. "Free your stuff Luxembourg", a Facebook group launched in 2011 to provide an online platform for exchanging consumer goods for free, has members with highly diverse national and linguistic backgrounds in the historically trilingual nation of Luxembourg, a small country of half a million; 44 % of residents are non-Luxembourgish nationals, and most people use more than one language in their everyday lives. The site's founder, a Romanian migrant, posted the first wall post in English, advertising a "plastic broom, green and grey". In the early phases, wall posts began in English, an increasingly common language in multilingual Luxembourg (in addition to Luxembourgish, the national language, French, and German), but replies followed in Luxembourgish and German (1).

(1) Nora: *I really need a bike!! Can somebody give me one or tell me where I can get a cheap one?*
Katy [German]: *Das hängt davon ab, was du unter "cheap" verstehst.*
'It depends what you mean by "cheap"'.
Anne [Luxembourgish]: *Kuck mol bei Luxbazar do gin der heiensdo verschärbelt oder wart op deen nexten Sperrmüll, do leien heiensdo nach gudder dobai!*
'Check Luxbazar, people sometimes give them away or wait for the next bulky waste disposal, there are sometimes quite good ones'
(adapted from Belling and de Bres 2014: 79–80).

Interestingly, these trilingual exchanges violate established practice in spoken interactions, in which Luxembourgers accommodate to their addressee's language. On this Facebook site, however, members respond in the language of their choice, assuming that addressees will be able to understand. Later, members actively debated language choices, with some opposing English. The group then adopted a policy encouraging members to write in the languages they knew best and suggested that persons needing or gifting items could request a translation if needed. Not everyone agreed, and disagreements erupted. Over time, use of English decreased substantially, and Luxembourgish increased, so that a once-diverse language community became nearly entirely monolingual.

By contrast, Seneweb, a New York-based web portal, caters largely to a recently-emigrated Senegalese diaspora living mainly in Europe and the United States, but also has substantial participation from Senegal's capital, Dakar. Stories on Seneweb are, without exception, written in French, Senegal's official language. In everyday life, however, Wolof is Senegal's dominant language, spoken by about 90 % of the multilingual-multiethnic population as a first or second language. Seneweb posts are predominantly in French, with a small minority in Wolof, which has no established written norms. In contrast to early multilingual exchanges on "Free your stuff Luxembourg", where everyone used one language at a time, some Seneweb speakers switch and mix words from French, English, and Wolof in the same utterance (2):

(2) Thieuy Wade, yakh na Sénégal, *il a rendu* niakk fayda *la fonction présidentielle* ba kou diokk reek di ladj rewmi. Ko khamné CE2 nga bayé *lékole* no meuné yorép rék, ko khamné *kartier* medina sakh meuno ko *gérer. Vraiment* **You** dafa yébaté *koi*. [French in italics, English in bold]
'Oh Wade, he's ruined Senegal, he's debased the office of President to the point where anyone can stand up and ask to run the country. How can someone who dropped out of school in CE2 do it, someone who couldn't even run the Medina neighborhood. You is really messing around.'
(adapted from McLaughlin 2014: 33)

4. Changing face of multilingualism and superdiversities old and new

The changing face of multilingualism in today's world is manifest in several trends, some already illustrated by examples of multilingual communities. These include new waves of transnational migration bringing new groups and languages to cities like London, and the introduction of new digital forms of communication like Seneweb facilitating global connections among diasporic communities. While some argue that these forces have made our world no longer simply 'diverse', but 'superdiverse', at the same time, the expansion of languages through European colonization of the New World and the more recent processes of globalization have spread a small handful of languages, especially English, on a scale never before possible. Ninety-five percent of the world's population speaks 5 % of its languages, while most of the world's languages (55 %) are spoken by a tiny fraction (less than 1 %) of the population. Even as language continues to be the companion of empire in the emergent post-colonial world order, its imposition derives less from military force and conquest than from economic control of world markets. The language Charles V supposedly reserved for birds has become the first truly global lingua franca, increasingly making English the most widely spoken language in human history and an indispensable part of multilingual, polyglossic repertoires. Two thousand years ago, there were only about 250 million people in the world; now, nearly 400 million speak English as a first language, and as many as two billion use it as a second (Svartvik and Leech 2016). Financial institutions and corporations in English-speaking countries have dominated world trade and made English the international language of business, publishing, and science. Although English dominates the internet, with circa 985 million users (Internet World Stats 2017), most users are nonnative speakers. English is also the dominant *de facto* or official language in over 90 countries. Neoliberal policies of economic restructuring imposing English on ever-more domains of global life have accelerated adoption of English-medium instruction, in many cases beginning in the early school years (Romaine 2016).

The idea that contemporary multilingualism is novel in its complexity and heterogeneity originated from outside linguistics in Vertovec's (2007: 1024) claim that "diversity in Britain is not what it used to be" due to changing patterns in global migration flows creating "a level and kind of complexity surpassing anything the country has previously experienced". London, as the predominant locus of immigration, is the place where superdiversity is "at its most marked" (Vertovec 2007: 1042). Scholars in various fields

have since adopted the term 'superdiversity' to describe a transformative 'diversification of diversity'. Aronin and Jessner (2015) distinguish between historical and current multilingualism, arguing that the period after the Second World War ushered in a new linguistic world order characterized by multilingualism as "the new linguistic dispensation", where every facet of life depends on multilingual social arrangements and individuals, whether directly or indirectly.

Some researchers propose superdiversity as the foundation for a sociolinguistics or applied linguistics more suited to the challenges of describing multilingual repertoires and transcending limitations of the predominantly monolingual orientation of modern linguistic theory that has framed multilingualism as exceptional, special, and aberrant (Blommaert 2015). New terms like 'translanguaging', 'metrolingualism', 'polylingualism', and 'lingua franca multilingualism' have been introduced to account for cases where languages are so deeply intertwined that it is difficult to determine boundaries between them (Garcia and Li Wei 2014). While these terms are new, the phenomena they describe are not. Not everyone endorses contemporary exceptionalism. This supposed paradigm shift away from essentialist notions of language and the boundedness of speech communities is primarily terminological (Flores and Lewis 2016; Pavlenko 2019). The idea that immigration has become more diverse may be partly a Eurocentric perspective (Czaika and de Haas 2014: 314). By comparison with other parts of the world, Europe was much less linguistically diverse to begin with, and many European nations were primarily sending migrants elsewhere rather than receiving them. Moreover, it is not clear when diversity becomes superdiversity. Why, for instance, do Belling and de Bres (2014) refer to the multilingual exchanges on "Free your stuff Luxembourg" as 'digital superdiversity', particularly when communication became less diverse over time, with most using Luxembourgish exclusively? Small-scale rural multilingualism arguably displays a degree and kind of complexity comparable to that of superdiverse cities like London, and urban multilingualism dates back to the earliest Mesopotamian city-states.

5. Conclusions and directions for future research

Encompassing a wide variety of complex sets of communicative practices that evolve differently in different settings, multilingualism is a fact of daily life for most of the world, rather than a recent phenomenon transforming formerly homogeneous places and communities into superdiverse spaces. The search for the ideal, balanced bilingual with native-like fluency is elusive, largely an artifact of a narrow theoretical perspective founded on monolingualism as its reference point. Given the diversity of contexts in which multilingualism exists, there is no such thing as a 'typical' multilingual family, country, community, or situation. Despite the 'multilingual turn' leading to a welcome, increased interest in multilingualism (May 2014), challenges remain in overcoming the hegemony of monolingualism and dominant language ideologies. Meeting these will require interdisciplinary collaboration and innovative methodologies (Martin-Jones and Martin 2017). Further research is needed in areas like assessment of multilingual competences (Treffers-Daller and Silva Corvalán 2016) and their evolution across the lifespan (Nicoladis and Montanari 2016), as well as multilingualism involving signed languages (Schembri and Lucas 2015).

6. References

Adams, James N.
2003 *Bilingualism and the Latin language*. Cambridge: Cambridge University Press.

Armstrong, Timothy
2012 Establishing new norms of language use: The circulation of linguistic ideology in three new Irish-language communities. *Language Policy* 11(2). 145–168.

Aronin, Larissa & Ulrike Jessner
2015 Methodology in bi- and multilingual studies: From simplification to complexity. *AILA Review* 27. 56–79.

Baker, Colin & Sylvia Prys Jones
1998 *Encyclopedia of bilingualism and bilingual education*. Clevedon: Multilingual Matters.

Belling, Luc & Julia de Bres
2014 Digital superdiversity in Luxembourg: The role of Luxembourgish in a multilingual Facebook group. *Discourse, Context and Media* 4–5. 74–86.

Blommaert, Jan
2015 Commentary: Superdiversity old and new. *Language and Communication* 44. 82–88.

Bloomfield, Leonard
1933 *Language*. Chicago: University of Chicago Press.

Bolton, Kingsley & Bee Chin Ng.
2014 The dynamics of multilingualism in contemporary Singapore. *World Englishes* 33(3). 307–318.

Burke, Peter
2004 *Languages and communities in early modern Europe*. Cambridge: Cambridge University Press.

Campbell, Lyle & Verónica Grondona
2010 Who speaks what to whom? Multilingualism and language choice in Misión La Paz. *Language in Society* 39(5). 617–646.

Cenoz, Jasone
2013 Defining multilingualism. *Annual Review of Applied Linguistics* 33. 3–18.

Chomsky, Noam
1965 *Aspects of the theory of syntax*. Cambridge: MIT Press.

Czaika, Mathias & Hein de Haas
2014 The globalization of migration: Has the world become more migratory? *International Migration Review* 48(2). 283–323.

de Nebrija, Antonio
1989 [1492] *Gramàtica de la lengua castellana*. [Grammar of the Castilian language]. Antonio Quilis (ed.). Madrid: Centro de Estudios Ramón Areces.

Diebold, A. Richard
1961 Incipient bilingualism. *Language* 37(1). 97–112.

Edwards, John
2007 Societal multilingualism: Reality, recognition and response. In Peter Auer & Li Wei (eds.), *Handbook of multilingualism and multilingual communication*, 447–469. Berlin: Mouton de Gruyter.

Evans, Nicholas
2012 Even more diverse than we had thought: The multiplicity of Trans-Fly languages. In Nicholas Evans & Marian Klamer (eds.), *Melanesian languages on the edge of Asia: Challenges for the 21st century. Language Documentation & Conservation Special Publication* 5, 109–149. Honolulu: University of Hawaii Press.

Fishman, Joshua A.
1965 Who speaks what language to whom and when? *La Linguistique* 1. 67–88.

Fishman, Joshua A.
1991 *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon: Multilingual Matters.
Flores, Nelson & Mark Lewis
2016 From truncated to sociopolitical emergence: A critique of super-diversity in sociolinguistics. *International Journal of the Sociology of Language* 241. 97–124.
García, Ofelia & Li Wei
2014 *Translanguaging: Language, bilingualism, and education*. New York: Palgrave Macmillan.
Grosjean, François
1997 Processing mixed language: Issues, findings, and models. In Annette M.B. de Groot & Judith F. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, 225–254. Mahwah: Lawrence Erlbaum.
Internet World Stats
2017 http://www.internetworldstats.com/stats7.htm (last accessed 9 November, 2017).
Martin-Jones, Marilyn & Deidre Martin (eds.)
2017 *Researching multilingualism: Critical and ethnographic perspectives*. London and New York: Routledge.
May, Stephen (ed.)
2014 *The multilingual turn: Implications for SLA, TESOL, and bilingual education*. New York: Routledge.
McLaughlin, Fiona
2014 Senegalese digital repertoires in superdiversity: A case study from Seneweb. *Discourse, Context and Media* 4–5. 29–37.
Mehmedbegović, Dina, Peter Skrandies, Nick Byrne & Philip Harding-Esch
2015 *Multilingualism in London*. London: London School of Economics Language Centre.
Nicoladis, Elena & Simona Montanari (eds.)
2016 *Bilingualism across the lifespan: Factors moderating language proficiency*. New York: American Psychological Association and De Gruyter Mouton.
Ostler, Nicholas
2005 *Empires of the word: A language history of the world*. New York: HarperCollins.
Parpola, Simo
2004 National and ethnic identity in the Neo-Assyrian empire and Assyrian identity in post-empire times. *Journal of Assyrian Academic Studies* 18(2). 5–22.
Pavlenko, Aneta
2019 Superdiversity and why it isn't: Reflections on terminological innovations and academic branding. In Barbara Schmenk, Stephan Breidbach & Lutz Küster & (eds.), *Sloganizations in language education discourse*. *Thinking in the age of academic marketization*. Bristol: Multilingual Matters.
Romaine, Suzanne
1995 *Bilingualism*, 2nd edn. Oxford: Blackwell.
Romaine, Suzanne
2016 Linguistic diversity and global English: The pushmi-pullyu of language policy and political economy. In Thomas Ricento (ed.), *Language policy and political economy: English in a global context*, 252–275. Oxford: Oxford University Press.
Schembri, Adam C. & Ceil Lucas (eds.)
2015 *Sociolinguistics and deaf communities*. Cambridge: Cambridge University Press.
Simons, Gary F. & Charles D. Fennig (eds.)
2017 *Ethnologue: Languages of the world*, 20th edn. Dallas, TX: SIL International. http://www.ethnologue.com.

Singer, Ruth & Salome Harris
2016 What practices and ideologies support small-scale multilingualism? A case study of Warruwi community, northern Australia. *International Journal of the Sociology of Language* 241. 163–208.
Singleton, David, Joshua A. Fishman, Larissa Aronin & Muiris Ó Laoire (eds.)
2013 *Current multilingualism: A new linguistic dispensation*. Berlin: De Gruyter Mouton.
Strevens, Peter
1982 The localized forms of English. In Braj Kachru (ed.), *The other tongue: English across cultures*, 23–30. Oxford: Pergamon.
Svartvik, Jan & Geoffrey Leech
2016 *English: One tongue, many voices*. London: Palgrave Macmillan.
Treffers-Daller, Jeanine & Carmen Silva Corvalán (eds.)
2016 *Language dominance in bilinguals: Issues of operationalization and measurement*. Cambridge: Cambridge University Press.
Vertovec, Steven
2007 Super-diversity and its implications. *Ethnic and Racial Studies* 30(6). 1024–1054.
Weinreich, Uriel
1968 [1953] *Languages in contact: Findings and problems*. The Hague: Mouton.
Wenger, Etienne
1998 *Communities of practice: Learning, meaning and identity*. Cambridge: Cambridge University Press.

Suzanne Romaine, Oxford (UK)

23. From geographical and social boundaries to epistemic breaks

1. The notion of boundary
2. Setting language boundaries in multilingual settings
3. Languages and their structural boundaries
4. Communication strategies and language usage
5. The role of the descriptor and the breaking point
6. References

1. The notion of boundary

Boundaries exist in the physical world, but just as often are constructed by human behavior. That is the case for geographical boundaries, past and present. Boundaries *only* exist because 'humans' existed before them, and because humans identify, posit, recognize, and manipulate them (Nicolaï 2017).

Various types of boundaries can be identified. Some structure space in ways we have little control over: physical boundaries, such as the doorstep of a house, the Atlantic, or

https://doi.org/10.1515/9783110435351-023

the Sahara. Others only exist conceptually, resulting from and having an impact on our social, economic, political, and cultural human activities, yet being just as prevalent in and part of our history.

In terms of intellectual activities and constructs, the social and cultural categories which we devise and are unique to our societies presuppose, on the one hand, an examination of their (designated) boundaries and the (designated) meanings associated with them. On the other, these activities and constructs (and the boundaries they entail) can be questioned, notwithstanding the fact that they can also be intangible markers of our environment and lifestyles.

In sum, mastering and delimiting boundaries are necessary for defining what is meaningful *for us*, since much of our activity is based on recognizing distinctions. In our daily lives, we spend time identifying boundaries, implementing them, calling them into question, and/or confirming them, i.e. examining their dynamics, their functionalities, their impact, the ways they (we) enable us (them) to (re)structure and understand the world around us. They are at once a constituent part and the object of both our ordinary activities and our intellectual constructs. We necessarily take them into account.

Is this true for language? Issues of linguistic boundaries, and phenomena linked to contact, provide relevant material. Indeed, languages are *inside us*, they necessarily exist in our minds individually and collectively, but they are also *outside us*, materialized in the utterances we exchange, omnipresent in the explicit and implicit content of our discourse, conversations, and literary output. They can be codified in dictionaries and grammars. They can also be ascribed to social groups and potentially attributed to geographical areas. Thus, languages as signifiers and communicative tools are systematically objectified, because we recognize them and endow them with stable forms and structures at various levels of understanding: we organize them into systems, and are able to interpret them through symbols. These activities all presuppose that their boundaries are already established (Nicolaï 2007, 2018; Nicolaï and Ploog 2013). That being said, forms, structures, and systems, on the one hand, and symbolic representations, on the other, exist only because of our own recognition of time-bound standards and requirements, and because we have prior theoretical and referential frameworks to help us understand and conceptualize them in their structural state and symbolic functionalities. They are thus both ubiquitous and continuously reinvented.

Despite this near-transcendence and the more or less stable standardized representations we have of them, languages are in constant renewal. At a pace that can be slow, but often perceptible, over a lifetime, they are continuously transformed. Independent of any experience of bi- or multilingualism, they are put ‘in contact’ with themselves through a reflective distancing process, backed by an inner boundary that conditions the use we make of them − unless we rely on a specious reduction which, on the one hand, would posit single communicative functionalities and strict representational transparency in our language practices and use and, on the other, would negate our metalinguistic and epilinguistic activities. Apart from this inherent reflexivity, languages are also perceived as homogeneous and countable entities, which legitimizes the practice of counting them (the total number of currently spoken languages is estimated between 3000 and 7000). Moreover, any language user (sociolinguist or not) can observe that languages are perceived as objects that can be manipulated and (neatly) classified. Thus, while there is always a level of understanding where those who use language call upon stabilized

representations considered to be references (with clear boundaries), there is also a level of understanding where languages are variable and made up of a transient mix, determined by social, stylistic, and adaptive functionalities (implying more porous and dynamic boundaries).

Languages are difficult objects − constructed, reconstructed, and functionalized depending on communicative needs, scientific knowledge, and our social and political environment. These objects are always recognized, never denied, and are considered obvious. They are predetermined social and cognitive artefacts, made and used by humans − roaming animals with a strong tendency to band together around a community structure, and to continuously redefine their space and rules of procedure, thereby creating and manipulating boundaries. Given the link between society, politics, and geography, it becomes possible to examine the effects of the social, political, geographical and other boundaries in languages and to study them from these viewpoints. I therefore devote some space to the notions of 'boundary' and 'contact', which, despite the apparent obviousness of our intuitive understanding, mask and retain a tremendous degree of complexity.

I first examine the social and political activities of multilingual communities, where the issue of linguistic boundaries can be significant. I then consider some of our linguistic practices when we interact in conversations, and develop strategies pertaining to linguistic boundaries which, implicitly or explicitly, we recognize, establish, or erase. I conclude by considering the perspective of descriptors as a working account for the observed facts; we shall see that the effects of establishing the boundaries they observe, and those which they introduce, have significant consequences.

2. Setting language boundaries in multilingual settings

All over the world, there are (political, social, cultural, and geographical) areas characterized by greater homogeneity than others. This may lead to the belief that within such areas, there is little or no linguistic diversity, and that therefore questions of boundaries would only affect their margins, given that, in correlation, bilingualism is not shared by all. However, if we change the scale, this is in fact an oversimplification, rather than an empirical observation. Before exploring this, consider Mandelbrot's image (1982: 17–18) to explain the notions of scale and dimensions underlying his approach to fractals:

> […] a ball of 10 cm diameter made of a thick thread of 1 mm diameter possesses (in latent fashion) several distinct effective dimensions. To an observer placed away, the ball appears as a zero-dimensional figure: a point. […] As seen from a distance of 10 cm resolution, the ball of thread is a three-dimensional figure. At 10 mm, it is a mess of one-dimensional thread. At 0.1 mm, each thread becomes a column and the whole becomes a three-dimensional figure again. At 0.01 mm, each column dissolves into fibers, and the ball again becomes one-dimensional, and so on, with the dimensional crossing over repeatedly from one value to another. When the ball is represented by a finite number of tomblike pinpoints, it becomes zero-dimensional again.

Each change in the 'degree of resolution' represents a change in viewpoint, a relevant new choice, and therefore a *discontinuity* in the process of understanding.

This image is transposable because, depending on the focus human actors adopt, their approaches to language will be incommensurate. Depending on the approach, they will identify, erase, or establish boundaries. The framework that a linguist uses while doing fieldwork is not what a politician uses to analyze political discourse, nor what speakers use to make sense of the language(s) used in their daily surroundings. Moreover, once the focus has been set, linguists, politicians, and ordinary speakers, depending on the theoretical, practical, ideological, and topical criteria they (un)consciously adhere to, will still be confronted with various, complex ways of understanding the linguistic object. The interaction between human actors, and the phenomena they are attempting to understand, is thus obvious. In what follows, I illustrate these changes in focus to distinguish 'areas of reference', and show the limits and effects of what I consider 'setting boundaries'. I concentrate on homogenizing globalizing approaches, on the one hand, and approaches that take heterogeneity at the level of states and regions into account, on the other.

2.1. Setting boundaries from a global perspective

An example of the apparently simple notion of *World Language* used in a seemingly objective way can be found in the globalizing approaches of international organizations such as the *Foundation for International Development Study and Research* (FERDI). In some FERDI publications, areas are presented as monolingual or multilingual from a geopolitical perspective. This makes it possible to convert monolingual and multilingual areas into variables in the calculations, forecasts, and recommendations that such organizations are expected to provide. The scale is global.

According to FERDI (www.ferdi.fr), a 'world language' includes all countries with over 500,000 inhabitants satisfying either of the following conditions: (i) the *de jure* dimension, i.e. the language in question is the official language; (ii) the *de facto* dimension, i.e. a significant portion (20 %) of the inhabitants speak the language in question – thus certain countries may belong to several areas at once (Carrère and Masood 2014). As such, FERDI not only tries to calculate the linguistic weight of a certain linguistic area, but also tries to match the linguistic data of an area with other available data (including population and economic data) concerning the same area. The results are over-simplifications. For example, according to FERDI criteria, the English-speaking area covers 43 countries, including: Australia, Botswana, Cameroon, Canada, Egypt, Eritrea, Fiji, Gambia, Ghana, Guyana, Hong Kong, India, Ireland, Israel, Jamaica, Jordan, Kenya, Kuwait, South Africa, the UK and the USA.

It is easy to uncover the linguistic, as well as the social and political, heterogeneity hidden underneath the homogenizing 'English speaking' label. The dimensions used to distinguish language areas by an organization whose aim is to promote a fuller understanding of international economic development, and the factors that influence it, are not at all commensurate with those that linguists would be inclined to use. 'Linguistic boundaries' are being defined on different bases. In the FERDI universe, we are forced to approach the heterogeneous linguistic reality from a homogenizing global perspective. Using Mandelbrot's (1982) image of a 'ball made of a thick thread', the focal depth that we are forced to apply at a great distance is 'zero-dimensional'.

2.2. Setting boundaries in countries and regions

Exchanging the global perspective for a perspective focusing on countries and regions, we can experience language areas, in general, and language boundaries, in particular, in greater detail. Linguistic situations − certainly where two or more natural languages are used by various parts of the population − can be conflictual, characterized by latent or manifest power struggles between communities. These struggles are not purely language-related, but also linked to demographic, political, cultural, historical, identity-based, communicational, social, economic, and/or other variables. These variables, and the collective representations they introduce (Anderson 1983), have an impact on the attitudes about the languages used in a given setting and the way in which the language and other boundaries in that very setting are defined. Attitudes have a strong influence on the way in which 'realities' are constructed and reconstructed (Barth 1969). In what follows, the variety of ways in which boundaries are determined in organized multilingual areas will be illustrated with the help of examples mainly taken from Lacorne and Judt (2002).

2.2.1. Setting boundaries out of the political desire for linguistic differentiation

In the former Yugoslavia, according to Garde (2002: 296), one not only notes

> *une intercompréhension totale, mais même une très grande homogénéité. Les différences dialectales existent certes, et sont un peu plus fortes aux marges, dans les zones de transition croato-slovène […] au Nord-Ouest, ou serbo-bulgaro-macédonienne au Sud-Est […]; mais elles sont moindres qu'à l'intérieur d'espaces beaucoup plus restreints comme la Slovénie ou la Suisse alémanique. Ajoutons que ces faibles différences ne coïncident pas avec les différences nationales (à base confessionnelle) entre Serbes, Croates, Bosniaques musulmans. Les habitants d'une même région parlent de même, les accents sont locaux, non nationaux.*
> [a total mutual understanding, but even a very large degree of homogeneity. Dialect differences do exist, and are more pronounced at the margins, in Croato-Slovenian transition areas […] in the Northwest, or in Serbo-Bulgaro-Macedonian transition areas in the Southeast […]; but they are less pronounced there than within much smaller areas such as Slovenia or German-speaking Switzerland. It should also be added that these slight differences do not coincide with (religion-based) national differences between Serbs, Croatians, and Bosnian Muslims. The inhabitants of a particular area speak in a similar way, accents are local, not national].

Technically, differences among varieties used in former Yugoslavia bear on choice of alphabet, orthography, accent, pronunciation of some phonemes, some points of morphology, and lexical differences. That said, the Croatians − as Garde (2002: 300−301) states − do sometimes experience how the unity of the 'Serbo-Croatian language' is being used as an argument to deny or minimize their national identity. Already in the 1920s and 1930s, the Croatians, in accordance with the principle 'one people, one language', increasingly started to stress the autonomy of the Croatian language as opposed to the Serbian. In response, Croatian was severely repressed during the first 20 years of Tito's reign, during which 'Serbo-Croatian' became dogma.

In this instance, language boundaries are not so much determined linguistically, but rather on the basis of extra-linguistic criteria used to accentuate sometimes minor features that distinguish the languages at the linguistic level, and as such help to more strongly symbolize the boundaries communities wish to establish (see Trudgill [2004] on the instrumentalization of linguistic features in modern processes of nation-building; see also Kloss [1967] on *Abstand*- and *Ausbau*-languages). Central here is the question raised by Busch and Kelly-Holmes (2004), whether languages cause borders or borders cause languages. Similar cases, albeit less linked to actual violence, include regional varieties of languages that are promoted as languages in their own right: Valencian as opposed to Catalan, Meänkieli as opposed to Finnish, etc.

2.2.2. Boundary maintenance to ascertain cultural pluralism in so-called melting pots

The United States was, in the past, described as a melting pot in which, over the years, differences between (immigrant) minority and majority communities would have largely melted away or fused into a harmonious whole. Fishman (1966) depicts the pattern of language shift as follows: the 1st generation of immigrants is monolingual, the 2nd is bilingual, and the 3rd becomes monolingual again in the 'new' language. However, language shift in many cases turned out to be anything but linear. In some areas of the American southwest, for centuries the dominant language was Spanish. A few years after the Gold Rush, the number of Spanish speakers had dropped to just 10 % of the population in California. In a few decades, some communities were largely absorbed by the English-speaking majority (Lopez 2002: 118). What counts for the descendants of Spanish speakers who arrived in the 19th century, however, does not seem to count any longer for the Spanish speakers who arrived in the course of the 20th. The chance that a 3rd generation descendant is bilingual is five times higher in the case of Latino communities than in the case of Asian communities (Lopez 2002: 127; Rumbaut, Massey, and Bean 2006). In fact, in Latino communities, language shift no longer seems to follow the traditional pattern identified by Fishman in the late 1960s. The reasons for this are manifold. They are related to strong community ties in large Latino communities and ongoing immigration continuously renewing the pool of Spanish-speakers, as well as the coexistence of linguistic and social boundaries against which movements, such as the Chicano Civil Rights movement, started to take action in the 1960s. Within the Latino community, the use of (varieties of) Spanish next to English has turned into an important symbol not meant to support separatist claims, but to consciously assert cultural pluralism in an ethnically diverse setting. This illustrates the importance of languages as symbolic boundary markers in the case of ethnic communities (or heritage language communities) that try to maintain their own language, while simultaneously trying to adapt to their environment (Potowski 2010).

2.2.3. The willingness to treat boundaries as subordinate to a political project of linguistic and cultural cohabitation

Another example of ways in which language boundaries are set is Switzerland, with its four national languages (German, French, Italian, and Romansh) of unequal size. What

keeps the country together is not so much the fact that a large number of Swiss people master two, three, or occasionally even four of their national languages. What keeps the country together is that the Swiss share a single political culture, that is, they are willing to support the project of a Swiss so-called *Willensnation*, i.e. a nation created by the will of members of different ethnic groups who do not share the same language, and who are also (even internally) marked by different religions (Grin 2014; Schläpfer et al. 1982).

According to Windisch (2002: 228), the coexistence of different cultural and linguistic communities as part of one, single nation-state depends on the simultaneous, joint presence of the following three components: (i) cultural identity, (ii) cross-cultural communication, and (iii) a political culture shared by all linguistic and cultural communities. If one of these components fails or starts to outweigh the others, then difficulties of a general political order start to appear that can only be resolved by taking all three previously mentioned components (identity, communication, and politics) into account and trying to restore equilibrium.

The list is long of situations where the (re)framing of language boundaries (either through downplaying or stressing their importance) color sometimes heated debates on the linguistic and cultural coexistence of different language communities, and interfere with top-down and/or bottom-up language policies in areas that are linguistically, geographically, culturally, and politically organized as states, nation-states, or (autonomous) regions within states (e.g. Watt and Llamas 2014).

2.2.4. Socially cherished language boundaries in the Grassfields in the 19th century

Beyond the context of nation-states, other areas illustrate another way of considering boundaries, e.g. the Grassfields region in Cameroon, where social and individual multilingualism traditionally were (and to a great extent still are) the norm. The Grassfields region is characterized by numerous language varieties which are closely related, constitute a branch of the Benue-Congo languages and have been in massive contact for centuries. It is an area where linguistic density and diversity are particularly high. A large proportion of the inhabitants (at least 50 %) speak at least two local languages perfectly. Largely taken from Warnier (1980), the description that follows describes a state − which has since disappeared − that corresponds to what was prevalent before the spread of Pidgin and European languages in the late 19th and 20th century, characterized by the use of languages without written form and by the absence of a regional language of communication.

Just prior to colonial conquest, the Grassfields counted circa 500,000 inhabitants organized in several hundred more or less sovereign chiefdoms, each holding sway over subjects numbering between 100 and tens of thousands. Almost all had a language distinct from that of the neighboring chiefdoms. These were often very closely related and mutually intelligible. The number of those who were not amounted to over fifty within an area the size of Brittany.

The chiefdoms lived in close contact with each other. Weekly markets brought together merchants and buyers speaking different languages. Given the absence of a regional language of communication, these relations depended on the multilingualism of

the individuals involved. In this context, multilingualism became diversified and multidirectional. It was driven by local economic specialization and trade. Moreover, commercial activities were grounded in friendship and matrimonial networks: 20 % to 50 % of married women in a chiefdom were born in another chiefdom, and virilocal residence was the norm (the bride went to live with her husband, accompanied by children).

A certain logic linked the social categories to each other. Merchants maintained formal friendship networks and married outside their community. Their wives came to join them. The children of a woman from another chiefdom learned her language and used it throughout their lives. Boys went to stay with their maternal uncle, and it was not uncommon for people to speak 3, 4, or 5 languages. Multilingualism was considered an achievement and a sign of personal development. At the same time, a parochial mindset was also prevalent and linguistic differences were part and parcel of the social and political identity of individuals. There was a conservative, not to say purist, attitude favoring linguistic differentiation, especially among elderly speakers who enjoyed an elevated status.

In this context, and without any question of political struggles, language boundaries appear to have been recognized, accepted, preserved, and perhaps even functionalized, which does not mean that there cannot also have been structural modifications due to language contact.

2.2.5. The dissolution of language boundaries in 'reconstituted' multilingual areas

Choosing yet another dimension of the 'ball made of a thick thread', I conclude this overview with a glimpse at two cases that illustrate how new contact varieties can emerge when large groups of − often young(er) − people mingle in dense urban settings (e.g. Nortier and Svendsen 2015). Once again, with its multilingualism, its superposition of European languages, and its strong urban growth, Africa offers interesting examples.

The first example is Iscamtho, a mixed language that emerged in the region of Soweto, a township near Johannesburg (Childs 1997), which is the result of contact between several South African languages (including Zulu as its matrix language, as well as Sotho, Afrikaans, English, etc.). Its lexicon is heterogeneous. Structurally, it largely reflects Bantu syntax and morphology. Reflecting the multilingual context of the black townships in South Africa, it is mostly spoken by young people, men, and marginalized people, but also by other South Africans. Today it appears in advertising, on the radio, and on television, and is spreading to neighboring countries. It is said to be the symbol of the new identity of the black community in the new South Africa.

Camfranglais is the second example (de Féral 2006), an urban Cameroonian language that came into being on the streets of Douala before spreading to all large cities in Cameroon, where it was preceded by Pidgin-English, a vehicular language highly popular in Cameroon in the 1970s and 1980s. This mixed language is grounded in French, English, Cameroonian Pidgin, and Cameroonian languages. Its lexicon is full of terms invented by the youth of Cameroon.

These emergent mixed languages are the result of highly complex patterns of language use in ethnically and linguistically heterogeneous settings, and illustrate how the contact-induced erosion of language boundaries (and the emergence of a 'reconstituted language' as a result) goes hand in hand with the creation of new urban identities.

In contrast to these examples, one may easily conclude that − depending on the political, social, historical, etc. context − the effects of language contact on populations can be perceived and reacted to in different ways. Linguistic boundaries can be strategically functionalized, and the changes to standard practice can be accepted or stigmatized.

3. Languages and their structural boundaries

Contact between different varieties obviously leads to linguistic change. Some examples are the emergence of koines in the case of dialect contact, with pidgins and creoles used in contexts of limited mutual intelligibility, and mixed languages, in what Meakins (2013) describes as situations of 'thorough bilingualism'.

A koine is a supra-local idiom used between speakers of different varieties within a dialect zone. This entails feature mixing, most often in the form of reductions and simplifications. Naturally, this raises the issue of assessing the prestige of the supra-local idiom: *a priori*, reference to a standard of communication is not especially prestigious, but that can change if socially valued elements are added. The dynamics that make koines emerge appear to be a specific case of the dynamics in the elaboration of vehicular languages (English, Hausa, Spanish, etc.), namely, the will to favor communication without rigidifying linguistic boundaries.

In all contexts where the absence of a shared language makes communication difficult, pidgins develop and creoles come into being. The first studies devoted to these languages date to the 19th century (see e.g. the works of Schuchardt; http://schuchardt.uni-graz.at). Creoles exist around the world, based on European (e.g. English, French, Spanish, Dutch, and Portuguese) and non-European languages (e.g. Arabic, Bantu, Ngbandi, and Malay). Research on pidgins and creoles contributes to opening reflections around the question of language boundaries, not only at the structural level, but also at the symbolic level, in the sense that the ties between pidgins and creoles, and the languages (as well as the language communities) they are historically linked to, need to be acknowledged (see also Chapter 9).

Mixed languages also provide insights into how language boundaries function (see also Chapters 2, 3 and 6). Two well-known examples are Michif and Media Lengua. Described by Bakker (1997), Michif is a Cree- and French-based mixed language spoken in Canada and the northern United States. Structurally, it is mostly composed of French nouns, Cree verbs, and lexical borrowings from other Amerindian languages. It is said to have emerged as an identity marker among Métis raised in both languages. Media Lengua, spoken in Ecuador (Quito region), has Spanish vocabulary (approximately 90 %) and Quechuan grammar. There are several theories on its origins. According to Muysken (1997), it emerged through the self-identification of native populations who identified neither with Quechuan nor with Spanish culture. The language appears to have resulted from a swift relexification process from Quechua to Spanish. The study of such languages only developed recently (Matras and Bakker 2008; Thomason 1997). This may be due to the strength of pre-existing theoretical frameworks, and therefore to considerations of *conceptual boundaries.*

4. Communication strategies and language usage

To continue with the image of a 'ball made of a thick thread', I reduce my focal distance even more. To do so, I point out certain sociolinguistic practices, and broach the topic of boundaries, not at the level of the recognition of languages transformed through contact, but rather in terms of their political management, their community appropriation, and the language practices induced by social and individual multilingualism. A few notions, referring to practices ranging from the most institutionalized in our social relations to the most idiosyncratic in our communication exchanges, will suffice.

Diglossia: this notion denotes the state of two linguistic varieties (Ferguson 1959) − or languages (Fishman 1967) − that coexist within a given territory and have distinct social and political statuses and functions, and where one is considered superior to the other. The social and linguistic boundary is recognized, and acted upon, in practice. We are here within the framework of social regulations where individuals do not show behavior which could be considered as conscious manipulation of language boundaries, but where everyday social linguistic practices nevertheless influence the way boundaries are perceived (language [variety] A is used in higher domains of language use, language [variety] B in lower domains) and where the variety of observed situations has furthered our understanding of how societal multilingualism functions, and how tensions between language communities can be explained (e.g. Kremnitz [1981] on 'linguistic conflicts'; Lüdi [1990] on 'polyglossia').

Code-switching: here, the involvement of individuals in the deliberate manipulation of language boundaries is effective, and can be observed at several levels. Not only is it manifested in communication strategies, where it manifests itself mainly at the lexical level (cf. Lingala-Français: *bagarçon wana bazalaka vantards trop*, '*ces garcons-là sont très vantards*', 'those boys over there are terrible swanks'), but it can also be studied from the point of view of the structural constraints and cognitive patterns susceptible to determining formal regularities in the actualization of its deployment. A great deal has been written on the subject, and is easily accessible (e.g. Gardner-Chloros 2009; Poplack 1980; Stell and Yakpo 2015; see also Chapter 14).

Focusing, crossing, footing, and dramatization: turning to the inner layer of the 'ball made of a thick thread', I have 'focusing' (Le Page and Tabouret-Keller 1985) as well as 'crossing' (Rampton 1997, 2005) in mind. Both notions are meant to designate activities, often of a fun nature, which could almost be called 'artistic performances'. Rampton (1997: 2) defines crossing as "switching into a language or dialect that isn't generally thought to belong to you. [… it] involves a distinct sense of movement across social or ethnic boundaries and whenever it occurs, it raises questions of legitimacy that in one way or another, people need to deal with − by joining in, by laughing, by studiously ignoring it, by commenting on it etc.". Finally, consider what I call *dramatization*, which is manifested in monolingual contexts (Nicolaï 2001) and found in Goffman's (1981) notion of *footing*. This is a distancing process that all speakers can introduce into their discourse, a way of strategically reinvesting discourse. This illustrates our reflexivity and the gap it institutes. It poses a boundary between the saying and what is said, *de facto* materializing contact within the language (contact between the language and its own self, as presented earlier). Crossing can be considered a specific instance of dramatization in multilingual contexts.

5. The role of the descriptor and the breaking point

I close by examining the focal point of analysts, and their interaction with the 'reality' of phenomena that they suggest taking into account. Let us begin by restating some facts:

- The multiplicity of languages in the world is a fact.
- Human communities are more often multilingual than monolingual, although individuals are more often monolingual when they belong to communities with a widespread and dominant language.
- Contact between human groups is natural; it exists and has always existed, so that language contact is the rule, rather than the exception.
- It is we, as actors of communication, language producers, and descriptors, but also as collective political and cultural entities, who intervene in setting language boundaries; the preceding text suggests some active limits to such dynamics.

Notions of language boundary are manipulated. At the level of users and communities: will we posit language 'purity', and if so, of what sort? Will we develop and promote this or that form? Will we take into account cultural, ethnic, or political aspects, and if so, how? Will we assert their homogeneity?

For each of these possibilities, their acknowledgment and a global decision on the state (and the manipulation) of the 'boundaries' of these languages are required. Which aspect should prevail: political, cultural, or economic? This *de facto* introduces a power struggle within the shared community space.

Academically, depending on whether we accept a theoretical framework which provides space for the heterogeneity of languages and their use, we must consider situations marked by structural lability, weakness, or variability in standards as legitimate objects of study. This includes reconstituted languages, language mixing, and massive borrowings; it involves linguistic areas, questions of code-switching, and dramatization effects. Finally, I suggest accounting for individual use, as speakers functionalize the codes they have at their disposal − and, at times, redefine them *ad hoc* − for communication.

In conclusion, broaching the notion of 'boundary', we realize that objectivity is never obvious, never given, but always construed − at least in part. This is one reason why the question of contact between languages as a topic can been seen as 'non-topical' in theoretical frameworks which do not take linguistic heterogeneity into account, illustrating what I define as an epistemic break; to be likened, naturally, to Bachelard's (1938: 15) "*rupture épistémologique* [epistemological rupture]", for whom "*[l]e réel n'est jamais 'ce qu'on pourrait croire' mais il est toujours 'ce qu'on aurait dû penser'* [reality is never 'what one would have thought' but is always 'what one should have thought']".

Acknowledgments

My thanks go to Margaret Dunham for her translation; my thanks also to Jeroen Darquennes and Joe Salmons for their very useful and constructive comments.

6. References

Anderson, Benedict
1983 *Imagined communities: Reflections on the origin and the spread of nationalism*. London & New York: Verso.

Bachelard, Gaston
1938 *La formation de l'esprit scientifique*. Paris: Vrin.

Bakker, Peter
1997 *A language of our own: The genesis of Michif, the mixed Cree-French language of the Canadian Métis*. Oxford: Oxford University Press.

Barth, Fredrik
1969 *Ethnic groups and boundaries: The social organization of culture difference*. Boston: Little, Brown and Company.

Busch, Brigitta & Helen Kelly-Holmes (eds.)
2004 *Language discourse and borders in the Yugoslav successor states.* Clevedon: Multilingual Matters.

Carrère, Céline & Maria Masood
2014 *Le poids économique des principaux espaces linguistiques dans le monde.* Ferdi. http://www.ferdi.fr/sites/www.ferdi.fr/files/publication/fichiers/rapport_espace_linguistique_imp_1.pdf (accessed 28 august 2018).

Childs, G. Tucker
1997 The status of Isicamtho, an Nguni-based urban variety of Soweto. In Arthur K. Spears & Donald Winford (eds.), *The structure and status of pidgins and creoles*, 1–27. Amsterdam: John Benjamins.

de Féral, Carole
2006 Décrire un "parler jeune": Le cas du camfranglais (Cameroun). *Le français en Afrique* 21. 257–265.

Ferguson, Charles A.
1959 Diglossia. *Word* 15. 325–340.

Fishman, Joshua A.
1966 *Language loyalty in the United States: The maintenance and perpetuation of non-English mother tongues by American ethnic and religious groups*. The Hague: Mouton.

Fishman, Joshua A.
1967 Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues* 23(2). 29–38.

Garde, Paul
2002 Unité et pluralité dans le domaine linguistique serbo-croate. In Denis Lacorne & Tony Judt (eds.), *La politique de Babel*, 295–319. Paris: Karthala.

Gardner-Chloros, Penelope
2009 *Code-switching.* Cambridge: Cambridge University Press.

Goffman, Erving
1981 *Forms of talk*. Philadelphia: University of Pennsylvania Press.

Grin, François
2014 *Le débat sur les langues en quinze questions: Arguments, faits et chiffres*. Genève: Observatoire Économie – Langues Formation.

Kloss, Heinz
1967 Abstand languages and Ausbau languages. *Anthropological Linguistics* 9(7). 29–41.

Kremnitz, Georg
1981 Du "bilinguisme" au "conflit linguistique": Cheminement de termes et de concepts. *Langages* 61. 63–74.

Lacorne, Denis & Tony Judt (eds.)
2002 *La politique de Babel: Du monolinguisme au plurilinguisme des peuples.* Paris: Karthala.

Le Page, Robert & Andrée Tabouret-Keller
1985 *Acts of identity*. Cambridge: Cambridge University Press.
Lopez, David
2002 Bilinguisme et changement ethnique en Californie. In Denis Lacorne & Tony Judt (eds.), *La politique de Babel*, 117–145. Paris: Karthala.
Lüdi, Georges
1990 Diglossie et polyglossie. In Günther Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der Romanistischen Linguistik* V/1, 307–334. Tübingen: Niemeyer.
Lüdi, Georges
1994 Qu'est-ce qu'une frontière linguistique. *Babylonia* 1, 6–17.
Mandelbrot, Benoît
1982 *The fractal geometry of nature*. New York: W. H. Freeman and Company.
Matras, Yaron & Peter Bakker (eds.)
2008 *The mixed language debate: Theoretical and empirical advances*. Berlin: de Gruyter.
Meakins, Felicity
2013 Mixed languages. In Peter Bakker & Matras Yaron (eds.), *Contact languages: A comprehensive guide*, 159–228. Berlin: de Gruyter.
Muysken, Pieter
1997 Media Lengua. In Sarah G. Thomason (ed.), *Contact languages: A wider perspective*, 365–426. Amsterdam: John Benjamins.
Nicolaï, Robert
2001 La "construction de l'unitaire" et le "sentiment de l'unité" dans la saisie du contact des langues. *Traverses* 2. 359–385.
Nicolaï, Robert
2007 *La vision des faits*. Paris: L'Harmattan.
Nicolaï, Robert
2017 *Signifier: Essai sur la mise en signification*. Lyon: ENS Édition.
Nicolaï, Robert
2018 Sous le dessin et le dessein, ou des conditions de la frontière et de l'émergence du sens. In Michelle Auzanneau & Luca Greco (eds.), *Dessiner les frontières*. Lyon: ENS Éditions. http://books.openedition.org/enseditions/8596. DOI: 10.4000/books.enseditions.8596.
Nicolaï, Robert & Katja Ploog
2013 Frontières. In Jacky Simonin & Sylvie Wharton (eds.), *Sociolinguistique du contact: Dictionnaire des termes et concepts*, 263–287. Lyon: ENS Édition.
Nortier, Jacomine & Bente A. Svendsen (eds.)
2015 *Language, youth and identity in the 21st century: Linguistic practices across urban spaces*. Cambridge: Cambridge University Press.
Poplack, Shana
1980 Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics* 18. 581–618.
Potowski, Kim (ed.)
2010 *Language diversity in the USA*. Cambridge: Cambridge University Press.
Rampton, Ben
1997 New ethnicities, liminality and interaction. *Working Papers in Urban Languages & Literacies*. King's College London.
Rampton, Ben
2005 *Crossing: Language and ethnicity among adolescents*. Manchester, UK & Northampton, MA: St. Jerome.
Rumbaut, Rubén G., Douglas S. Massey & Frank D. Bean
2006 Linguistic life expectancies: Immigrant language retention in Southern California. *Population and Development Review* 32(3). 447–460.

Schläpfer, Robert, Jachen C. Arquint, Iso Camartin, Walter Haas, Pierre Knecht, Ottavio Lurati & Florentin Lutz
1982 *Die viersprachige Schweiz*. Zürich: Benziger.
Stell, Gerald & Kofi Yakpo (eds.)
2015 *Code-Switching between structural and sociolinguistic perspectives*. Berlin: de Gruyter.
Thomason, Sarah G. (ed.)
1997 *Contact languages: A wider perspective*. Amsterdam: John Benjamins.
Trudgill, Peter
2004 Glocalisation and the Ausbau sociolinguistics of modern Europe. In Anna Duszak & Urszula Okulska (eds.), *Speaking from the margin: Global English from a European perspective*, 35–49. Frankfurt: Peter Lang.
Warnier, Jean-Pierre
1980 Des précurseurs de l'École Berlitz: Le multilinguisme dans les Grassfields du Cameroun au 19e siècle. In Alain Bouquiaux, Larry Hyman & Jan Voorhoeve (eds.), *L'expansion bantoue: Actes du Colloque International du CNRS, Viviers France, 4–7 avril 1977*, 827–844. Paris: SELAF.
Watt, Dominic & Carmen Llamas
2014 *Language, borders and identity*. Edinburgh: Edinburgh University Press.
Windisch, Uli
2002 Multiculturalisme et plurilinguisme: Le cas suisse. In Denis Lacorne & Tony Judt (eds.), *La politique de Babel*, 227–259. Paris: Karthala.

Robert Nicolaï, Nice (France)

24. Language and identity in language contact settings

1. Introduction
2. Ethnic identity as an object of study
3. Sociolinguistic approaches to ethnicity
4. Phonetic variation as a manifestation of identity choices of ethnic minorities
5. Methodological recommendations for future research
6. Conclusion
7. References

1. Introduction

Ethnicity has long been recognized as a potential "sociolinguistic boundary" (Rickford 1985) that simultaneously signals in-group affiliation and out-group differentiation. Indeed, ethnic groups often distinguish themselves culturally, including linguistically, from other groups with which they come into contact. As a result, two individuals who grew up in the same dialect region, or even in the same neighborhood, may sound quite different from each other. The development of distinctive ways of speaking, well documented for African Americans (Fasold 1972; Wolfram 1969, 2007), Louisiana Cajuns (Dubois and Horvath 1998, 2000), and Latinos (Fought 2003; Mendoza-Denton 2008;

https://doi.org/10.1515/9783110435351-024

Zentella 1990), has been taken to reflect and assert these groups' strong ethnic identities. Ethnic identity can manifest in various non-verbal ways, such as through closed social networks or strong ties to the homeland. The range of methodologies used to study ethnicity is presented in Section 3.2.

Despite wide recognition of the role of ethnic identity in defining social groupings, ethnicity has been elusive to operationalize, largely because of its inherent fluidity and intersectionality with other types of identity related to gender, class, and ostensibly, race. Theoretical and empirical treatments of ethnicity have overlapped significantly with those of race. While race has become widely recognized as a social construct created by and imposed on members of a community based on phenotypic qualities (Alim, Rickford, and Ball 2016), ethnicity is seen as the cultural reflection of a community's self-identification. Even though the separation of race and ethnicity along the lines of biology and culture, akin to definitions of sex and gender, has a solid theoretical foundation, drawing a clear line between ethnicity and race continues to pose challenges. First, researchers have historically conflated the two concepts, which makes the task of (re-)interpreting their findings problematic. Second, even research centered on ethnicity, rather than race, has had a tendency to assume that a given ethnic identity is true of an entire community, without paying sufficient attention to the different identities that members of these communities may hold, or the different expressions that these identities may take. In this chapter, I focus on ethnicity, but wherever appropriate, I highlight the ways in which ethnicity is embedded within the social construct of race.

Our increasingly nuanced understanding of ethnicity has necessitated shifts in how ethnic identity is conceptualized and studied. In this chapter, I discuss the move away from the traditional conceptualization of the speech of ethnic groups as static and bounded 'ethnolects', and toward a more agentive approach embodied in the 'ethnolinguistic repertoire'. I then outline a spectrum of methodologies that have been developed to capture the fluidity and complexity of ethnic identities. Next, I present key generalizations from studies that have investigated phonetic variation as a manifestation of identity in language contact situations. Lastly, I make recommendations for the direction of future research based on the limitations of existing studies. I argue that, to make real headway in our understanding of language and identity in increasingly more diverse language contact situations, future research must foreground the interconnectedness of ethnicity with other identities, and study ethnic groups in relation to each other rather than in isolation.

2. Ethnic identity as an object of study

2.1. Ethnicity is inherently fluid

The challenges with studying ethnicity begin with its definition. Ethnicity is broadly taken to reflect speakers' cultural heritage, but its substantial overlap with other social constructs, particularly that of race, defies neat classification. For example, many studies classify speakers as white, black, Latino, or Asian (Bell 1997; Bernstein 1993; Rickford 1985) based on phenotype. However, *a priori* ethnoracial classifications reflect social constructs built around power relations rather than speakers' self-identification. For instance, in 19th century America, the ethnoracial classification of European immigrants

was controversial, as 'whiteness' represented not just skin color, but also social privilege. Slavs, Italians, and Irish were seen as representing a different race than northern and western Europeans, on the grounds of characterological traits (the former were portrayed as uncivilized boors unfit for assimilation to American society) and skin color, variably labeled as 'not-yet-white', 'situationally white', 'not quite white', 'off-white', or 'conditionally white' (Roediger 2005). The long process of the 'whitening' of eastern and southern Europeans entailed a degree of cultural assimilation and adoption of existing US racial relations. These 'new immigrants' occupied a space of racial 'in-betweenness' for approximately half a century before being branded as 'white ethnics' in the post-World War II United States (Roediger 2005).

Other social factors that complicate definitions of ethnicity include the language spoken at home and religion. In the United States, a language-based classification system is often employed in academic research and in various community surveys. As a result, Spanish-speaking populations are often grouped under a common, all-encompassing label (such as 'Latino' or 'Hispanic'). Even though language is an indicator of cultural proximity, the use of a single label for speakers from a range of backgrounds assumes homogeneity where it may be lacking. In fact, few would claim that Mexicans in California and Puerto Ricans in New York are not culturally or linguistically distinct. Religion is another social criterion that complicates definitions of ethnicity. For example, it is primarily based on religious practice that sociolinguistic studies have treated Jews separately from other white Americans (Benor 2001; Fader 2009; Labov 1966; Laferriere 1979). Jews in America have been found to display regional linguistic variables at elevated rates (Labov 1966; Laferriere 1979), and to develop distinctively Jewish features, such as a released /t/ (Benor 2001; Levon 2006). However, the influence of religion on the cultural identity of other white ethnics, such as the overwhelmingly Catholic Italians and Poles, remains understudied.

Attempts at defining ethnicity based on any particular trait are problematic, since people's identities do not fit neatly into binary oppositions, such as 'ethnic' or 'mainstream', 'black' or 'white'. Rather, identity has been argued to be *multivalent* (Mendoza-Denton 2002): speakers simultaneously navigate a set of identifications that transcend ethnicity and intersect with other social categories, such as gender or class. Those identities may superficially seem contradictory, in that Mexican Americans may see themselves as 'Mexican' as well as 'American', and may choose to foreground one of these identities depending on the situation. In addition to these macro-level demographic categories, speakers' identities encompass micro-level distinctions that may only be relevant in a local context, such as the social category of a 'nerd' in American high schools (Bucholtz 2001). Layered identifications are particularly evident in multicultural or immigrant contexts, where local, national, and transnational norms may conflict.

Ethnicity is also hard to define because it is fluid. Ethnic affiliation may change over a lifetime, reflecting shifting identifications during childhood, adolescence, and adulthood. Similarly, the need to assert one's ethnic affiliation may change over time. For example, two African American women were shown to have reduced their use of African American English features after they joined the workforce (Rickford and Price 2013). The fluidity of ethnicity is also demonstrated by speakers who 'cross' ethnic boundaries by adopting speech features that project a different ethnicity. For example, white hip hoppers may adopt features of African American English, despite having no ties to African Americans (Cutler 1999). Similarly, in European multicultural contexts,

minority youth have been shown to take on features of each other's accents as a way of displaying affiliation (Rampton 1995).

2.2. Ethnicity intersects with other identities

Ethnic identity is co-constructed with other social identifications that speakers may adopt, such as those related to gender and place. The intersection of ethnicity and gender is well documented. For example, TH-stopping (the replacement of interdental fricatives with alveolar stops, as in [d]*is* for *this*) has been argued to be an ethnolinguistic variable in Cajun English, and yet it exhibits a gender effect, with Cajun men displaying an overall preference for the use of stops over fricatives (Dubois and Horvath 1998).

Ethnic identity is also closely interwoven with place identity. A minority speaker's adoption of the regional features of a local dialect is often taken to reflect cultural affiliation with, and linguistic assimilation to, the majority population. Hence, the acquisition of commonly recognized features of a given dialect region has provided the backbone for much of the linguistic literature on ethnic identity (see Section 4.2.). Yet, the adoption of regional variables may express an attachment to a given place, rather than to a given (majority) population. Regional features can be thought of as indexical of a geographic territory, whether it be a city, such as Pittsburgh (Johnstone and Kiesling 2008), or a neighborhood, such as New York City's Lower East Side (Becker 2010). This indexical link is even more pronounced in locales whose imagined 'ownership' has been contested, as in Labov's classic Martha's Vineyard study (Labov 1962). In his study, the realization of two diphthongs was predicted by ethnicity, with the English-descent group outperforming Portuguese- and Indian-descent groups. Labov interpreted the English-descent group's motivation to retain the island's centralized diphthongs as reflecting a strong place identity, supported by the fishermen's negative stance toward tourists.

3. Sociolinguistic approaches to ethnicity

3.1. Conceptual models of ethnic varieties

As our understanding of the complex nature of ethnicity has become more nuanced, the conceptual models of ethnic varieties have shifted accordingly. The term 'ethnolect' was traditionally used to differentiate the speech of ethnic groups from regional dialects. Initially, this term referred to the second language of bilingual immigrant groups that manifested interference from the first language, but gradually came to refer to differences in the way an ethnic group spoke the majority language (Becker 2014; Carlock and Wölck 1981). In the United States, varieties commonly described as ethnolects include African American English, Chicano English, and Cajun English. Speakers of these varieties share a common cultural background and exhibit systematic linguistic behavior that is attributable to a strong ethnic identity. In Europe, certain ethnolects have been argued to convey a common, minority status for speakers with a variety of ethnic backgrounds, and have thus been termed 'multi-ethnolects' (Clyne 2000: 87). This designation has

been applied to urban youth language in Sweden (Kotsinas 1988), Germany (Androutsopolous 2001), the Netherlands (Hinskens 2011; Nortier and Dorleijn 2008), Finland (Quist 2008), and England (Cheshire et al. 2011).

From a methodological standpoint, the term '(multi-)ethnolect' is useful in that it distinguishes the speech of a group that shares a common cultural background. However, the use of this term triggers assumptions about the linguistic and social coherence of ethnic groups that may not be valid. Positing the existence of an ethnolect assumes that members of a given ethnic group share a common grammar, an assumption that has rarely been tested empirically (Guy 2013). The ethnolect approach further assumes that all speakers of an ethnic variety manifest a given ethnic identity, and to the same degree. These two assumptions go against the complex and fluid nature of identity, and against empirical findings. Quantitative analyses of the speech of ethnic groups have often revealed vast variability at the group and individual level that needs to be better understood. For example, Clyne, Eisikovits, and Tollfree (2002) found that most speakers of Greek Australian English displayed subsets of phonological features identified as constitutive of this variety, but the subsets were largely overlapping. Similarly, Polish New Yorkers exhibit variability with respect to the use of several dialect features of New York City English, such as TH-stopping, a low back contrast, and a nasal split for short-*a*. However, these speakers' feature choice is driven by their symbolic orientation toward Poland, America, or the Polish community in New York City (Newlin-Łukowicz 2015, 2016). Thus, the superficial linguistic heterogeneity of this ethnic group in fact betrays the construction of different *types* of ethnic identities by Polish New Yorkers.

Contrasting with the idea of a static ethnolect is the notion of an "ethnolinguistic repertoire", defined as "a fluid set of linguistic resources that members of an ethnic group may use variably as they index their ethnic identities" (Benor 2010: 159). This view acknowledges the multiplicity of identities, ethnic and otherwise, that speakers may adopt, and allows for identities to fluctuate over time. Benor (2010) posits that speakers have access to and productively employ 'distinctive linguistic features' that are associated with various ethnic identities, not just their own. According to this view, a speaker's use of a given variable, such as TH-stopping, reflects their active construction of social categories. That is, sociolinguistic variables are argued to be linked to an "indexical field", defined as a "constellation of ideologically related meanings" (Eckert 2008a: 454) that inform speakers' linguistic choices in the creation of 'personas' or 'styles'. For example, Eckert (2008b) contends that in different communities, TH-stopping has come to reflect local meanings that have emerged from ideologies surrounding a given immigrant group. Thus, TH-stopping carries connotations of 'hard work' for Germans in Wisconsin (Rose 2006) and of 'toughness' for Mexicans in California (Mendoza-Denton 2008), and is associated with the prestige of the Cajun Renaissance in Louisiana (Dubois and Horvath 1998). As a result, the ethnolinguistic repertoire approach has the flexibility to account for the range of intra- and inter-speaker variation that is reported for ethnic groups, including the presence of speakers who go against established community norms.

Beyond ethnolinguistic studies, repertoire approaches to language have been applied to broader sociolinguistic phenomena, particularly to account for speakers' language choices in the context of globalization and the resulting super-diversity (Blommaert 2007; Blommaert and Backus 2013). For example, Blommaert (2007) applied the concept of a spatio-temporal scale, drawn from social theory, to analyze discourse phenomena as operating along vertical, rather than the traditionally horizontal, dimensions. The

scale model lends itself to discussions of ethnicity: the expression of one or the other type of identity could be interpreted as invoking different scale levels (e.g. national, local, and transnational).

3.2. Methodologies for studying ethnicity

Studies of identity in language contact situations have struggled to capture the complexity and fluidity associated with ethnicity. Just as there are many definitions and conceptualizations of ethnicity, there are as many ways to study it, ranging from qualitative approaches focused on ethnography and interaction (Harris 2006; Mendoza-Denton 2008) to quantitative methods. The latter have aimed to correlate speakers' rates of variable use with their ethnic identifications, broadly construed. Some researchers, like Dubois and Horvath (2000), have focused on speakers' social networks. Others have quantified speakers' ethnic orientation, understood as the "attitude toward a language and its speakers" (Nagy, Chociej, and Hoffman 2012: 1). Research in Toronto has shown that ethnic groups with a positive attitude toward their heritage culture assimilate less readily to the mainstream dialect than ethnic groups that do not share this attitude (e.g. Hoffman and Walker 2010; Nagy, Chociej, and Hoffman 2012; Nagy and Kochetov 2013). In a similar vein, Wong (2007) defined the Chinese American lifestyles of American-born Chinese using quantitative measures, and showed that those who engaged in activities associated with Chinese culture linguistically assimilated less to the surrounding dialect. Lastly, transnational ties to the homeland were found to determine which accent features were maintained in the speech of British Asians (Sharma 2014) and Polish New Yorkers (Newlin-Łukowicz 2015, 2016).

This methodological range highlights that ethnicity can be expressed through social networks, ethnic orientation, ethnic lifestyle, and transnational ties. Yet, these factors are likely to be correlated with each other, and are also challenging to separate from language use. Speakers with a lifestyle centered on their heritage culture are likely to put themselves in more situations where the heritage language is spoken. Similarly, if a speaker's social network is ethnocentric, the speaker is likely to display a positive attitude toward their heritage culture and its speakers.

4. Phonetic variation as a manifestation of identity choices of ethnic minorities

4.1. Development of distinct ethnolinguistic variables

The speech of ethnic minorities may differ from that of the majority population in the presence of distinct features that identify the speaker as unmistakably 'Mexican', 'Cajun', or 'Jewish'. In the case of bilingual groups, ethnolinguistic variables, be it in the domain of phonology, syntax, or morphology, typically originate as interference from the first language (L1). As these L1-influenced features become associated with social meanings, they develop into sociolinguistic variables, and are acquired as 'ethnic mark-

ers' by subsequent generations. When this happens, L1-influenced features are retained, despite the fact that the second generation may not speak the heritage language. The presence of ethnic-identified features often contributes to outsiders' perception of a 'foreign accent' in the speech of ethnic groups.

Chicano English provides a good illustration of the distinctiveness of ethnic varieties. This variety of English, spoken mostly in the American Southwest, is characterized by systematic influence from Mexican Spanish phonology and syntax. For example, in the domain of phonology, Chicano English speakers are believed to under-employ vowel reduction, monophthongize diphthongs, neutralize the /i/-/ɪ/ contrast, and produce stops for interdental fricatives (Santa Ana 1991). Although Chicano English exhibits Spanish-influenced linguistic features, it differs from the English spoken by second language learners whose first language is Spanish (Santa Ana 1993). First, it exhibits features that are absent from non-native speech (e.g. centralization of [ʊ]), and conversely, it does not manifest certain common features of a Spanish accent, such as the confusion of [ʃ] with [ʧ] (Fought 2003). Second, the degree of Spanish influence is believed not to recede with new generations of Chicano English speakers. Finally, Chicano English displays characteristics that seem to have originated in the speech of other non-white ethnic groups, such as African Americans (Fought 2003).

4.2. Selective adoption of regional dialect change

Ethnic groups have been shown to be selective in their adoption of regional dialect variables, which means that they may not sound like 'typical' speakers of the dialect. For example, Hispanics were reported not to accommodate to /aɪ/ monophthongization in North Carolina (Wolfram, Carter, and Moriello 2004). Similarly, the speech of African Americans and Mexican Americans in the American Midwest was marked by the absence or partial absence of the Northern Cities Shift (Gordon 2000; Konopka and Pierrehumbert 2008; Preston 2003). These two sound changes are salient characteristics of the local dialects, and in fact, define the South and Inland North as dialect regions (Labov, Ash, and Boberg 2006). Thus, their non-adoption by ethnic groups in these areas is striking, and has been attributed to ethnic orientation. For example, Preston (2003) explains the absence of the Northern Cities Shift in the speech of African Americans by the speakers' association of this sound change with the majority white population.

Ethnic groups do not always shun regional speech features. Studies covering diverse and metropolitan parts of the United States have reported that ethnic groups do participate in regional sound change, for example, in California (see Hall-Lew [2009] for Asian Americans), New York (see Wong [2007] for Asian Americans; Coggshall and Becker [2010] for African Americans), and Philadelphia (see Poplack [1978] for Puerto Ricans). Similar findings have been reported for Toronto (see Hoffman and Walker [2010] for Chinese and Italian Canadians) and London (see Sharma and Sankaran [2011] for British Asians).

The selective adoption of regional change has been linked to issues in second language acquisition. Specifically, the native-like acquisition of some regional features could be hindered by the speakers' L1 or the 'critical period'. This hypothesis could explain Koreans' and Russians' incomplete acquisition of the phonologically conditioned raising of /æ/ (i.e. the so-called short-*a* system) in Philadelphia (Friesner and Dinkin

2006; Lee 2000). Because short-*a* exhibits numerous lexical exceptions, it has been suggested that the conditions on raising can only be faithfully acquired by speakers whose families have been in the area for generations (Payne 1976; Roberts and Labov 1995).

The link between the acquisition of regional features and L2 competence also holds in cross-generational comparisons. Typically, later generations acquire local speech patterns more faithfully than the first generations do. For example, Hoffman and Walker (2010) found the linguistic conditioning for t/d-deletion to be more faithfully acquired by the second and third generations of Chinese and Italian immigrants (relative to the first generations). In fact, it has often been assumed that the first generation does not acquire local speech norms, but recent scholarship has suggested that a more nuanced view is appropriate. For example, Friesner and Dinkin (2006) have shown that Russians who moved to Philadelphia as teens acquired phonological norms of the dialect, even though they spoke with a Russian accent. They argued, however, that the variables were acquired in an order that mirrored their complexity, with short-*a* acquired last. Likewise, Polish immigrants in the United Kingdom have been shown to variably acquire local variants, such as glottal /t/ (Drummond 2011), the lack of the STRUT/FOOT distinction (Drummond 2013), and the alveolar variant of *-ing* (Drummond 2012; Schleef, Meyerhoff, and Clark 2011). The acquisition of these features is arguably linked to the immigrants' identities. For example, whether or not Polish immigrants plan to return to their home country in the future affects their realizations of *-ing,* with the standard velar form "signaling allegiance to the L1 identity" (Drummond 2012: 107), and the alveolar variant correlating with a desire to assimilate to British society.

5. Methodological recommendations for future research

The existing body of research on the interaction of ethnicity and language has produced some exciting findings that, by and large, suggest that ethnic identity shapes the way we speak. Because identity is complex and ever-changing, sociolinguists have been hard-pressed to develop the right methodologies for studying it. In this section, I argue that, in order to account for nuances in ethnic identity, we must study within-group differences in the speech of ethnic groups and analyze ethnicity in conjunction with other aspects of identity.

Existing sociolinguistic studies of ethnicity have overwhelmingly focused on the speech of the largest, non-white minority groups, and have rarely investigated within-group variation. For example, perception studies have suggested that African American, Asian American, Chicano, and Caucasian guises and voices are perceived to sound different from each other (e.g. Newman and Wu 2011; Purnell, Idsardi, and Baugh 1999), but linguistic differences beyond those broad categories have been assumed to be subtler, or even nonexistent. No matter how subtle, within-group variation needs to be captured, as it can uncover social nuances that are critical for a full understanding of the construction of ethnic identity.

Even though speakers who are phenotypically classified as belonging to one category ('white' or 'black') often adopt the same set of sociolinguistic variables, the rates of use of these variables within these broad categories may differ. For instance, African Americans and Caribbean Americans in New York City exhibit a similar linguistic repertoire,

but nonetheless display systematic quantitative differences in their realization of regional variables (Blake and Shousterman 2010; Blake, Shousterman, and Newlin-Łukowicz 2015). While both groups are /r/-ful, Caribbean Americans have higher rates of /r/-fulness than African Americans. Moreover, while both groups show the tensing and raising of /ɔ/ typically associated with New York City, they produce the off-glide differently. Finally, while both group's realization of /oʊ/ is closer to a New York realization than to Caribbean Creole English varieties, off-glide differences exist between the two groups, with Caribbean Americans producing a more monophthongal variant that suggests influence from Caribbean Creoles. These subtle linguistic differences reveal social differentiation that exists between these two groups of black New Yorkers, but is not immediately apparent.

The assumption that little variation can be expected within major ethnoracial groups is also visible in the study of white speakers. White speakers are often seen as 'lacking' ethnicity, and as a result, serve as the 'unmarked' comparison group for the study of other ethnicities (Trechter and Bucholtz 2001). Contributing to this perception is the (now contested) hypothesis that white minorities assimilate to the mainstream society within the classic three generation model, forming a single, white variety perceived to be homogeneous (Becker and Coggshall 2009). As a result, even work that is careful to account for speakers' ethnicities has tended to collapse all white speakers together, perhaps with the exception of Jewish Americans. Yet there are plenty of indications that white ethnics warrant detailed treatment. For instance, in New York City, Polish Americans have been shown to adopt different regional dialect features to express various ethnic orientations (Newlin-Łukowicz 2015, 2016). A careful linguistic comparison of a range of ethnic groups residing in the same dialect area could illuminate the way these groups navigate their complex identities, especially in diverse contexts, where it is particularly evident that identities are co-created in relation to each other.

The sociolinguistic study of ethnoracial groups not only needs to explore within-group variation, but it also needs to broaden the scope of inquiry to include multiple aspects of speakers' identities, as these other types of identities are precisely the factors that will determine further social differentiation within ethnic groups. What it means to be a Polish New Yorker is intricately tied to, for instance, what it means to be a man or a woman in New York City of a particular age, religion, sexual orientation, etc.

6. Conclusion

This chapter provides an overview of existing research on the linguistic footprint of ethnic identity. Overwhelmingly, existing scholarship underscores the role of language as a rich resource from which ethnic groups draw to culturally differentiate themselves. As studies of ethnic and other types of identity have proliferated, our understanding of ethnicity has become more nuanced and our methodologies more varied, in line with ethnicity's complex and fluid nature. In order to further advance the study of ethnicity in language contact situations, sociolinguists need to explore within-group variation, and to account for other types of identities that speakers project through language.

7. References

Alim, H. Samy, John R. Rickford & Arnetha F. Ball
2016 *Raciolinguistics*. New York: Oxford University Press.

Androutsopolous, Jannis
2001 From the streets to the screens and back again: On the mediated diffusion of ethnolectal patterns in contemporary German. *LAUD Linguistic Agency* A522.

Becker, Kara
2010 *Social conflict and social practice on the Lower East Side: A study of regional dialect features in New York City English*. New York: New York University, PhD dissertation.

Becker, Kara
2014 Linguistic repertoire and ethnic identity in New York City. *Language & Communication* 35(1). 43–54.

Becker, Kara & Elizabeth Coggshall
2009 The sociolinguistics of ethnicity in New York City. *Language and Linguistics Compass* 3(3). 751–766.

Bell, Allan
1997 The phonetics of fish and chips in New Zealand: Marking national and ethnic identities. *English World-Wide* 18(2). 243–270.

Benor, Sarah Bunin
2001 The learned /t/: Phonological variation in Orthodox Jewish English. *University of Pennsylvania Working Papers in Linguistics* 7(3). 1–16.

Benor, Sarah Bunin
2010 Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity. *Journal of Sociolinguistics* 14(2). 159–183.

Bernstein, Cynthia
1993 Measuring social causes of phonological variation in Texas. *American Speech* 68(3). 227–240.

Blake, Renee & Cara Shousterman
2010 Second generation West Indian Americans and English in New York City. *English Today* 26(3). 35–43.

Blake, Renee, Cara Shousterman & Luiza Newlin-Łukowicz
2015 African American language in New York City. In Sonja Lanehart, Lisa Green & Jennifer Bloomquist (eds.), *Oxford handbook of African American language*, 280–298. Oxford: Oxford University Press.

Blommaert, Jan
2007 Sociolinguistic scales. *Intercultural Pragmatics* 4(1).1–19.

Blommaert, Jan & Ad Backus
2013 Superdiverse repertoires and the individual. In Ingrid de Saint-Georges & Jean-Jacques Weber (eds.), *Multilingualism and multimodality: Current challenges for educational studies*, 11–32. Rotterdam: Sense Publishers.

Bucholtz, Mary
2001 The whiteness of nerds: Superstandard English and racial markedness. *Journal of Linguistic Anthropology* 11(1). 84–100.

Carlock, Elizabeth & Wolfgang Wölck
1981 A method for isolating diagnostic linguistic variables: The Buffalo ethnolects experiment. In David Sankoff & Henrietta Cedergren (eds.), *Variation Omnibus*, 17–24. Edmonton: Linguistic Research.

Cheshire, Jenny, Paul Kerswill, Sue Fox & Eivind Torgersen
2011 Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics* 15(2). 151–196.

Clyne, Michael
2000 Lingua Franca and ethnolects in Europe and beyond. *Sociolinguistica* 14. 83–89.
Clyne, Michael, Edina Eisikovits & Laura Tollfree
2002 Ethnolects as in-group varieties. In Anna Duszak (ed.), *Us and others: Social identities across languages, discourses and cultures*, 133–157. Philadelphia: John Benjamins.
Coggshall, Elizabeth & Kara Becker
2010 The vowel phonologies of African American and white New York City residents. In Malcah Yaeger-Dror & Eric Thomas (eds.), *African American English speakers and their participation in local sound changes: A comparative study*, 101–128. Raleigh: Duke University Press.
Cutler, Cecilia A.
1999 Yorkville crossing: A case study of hip hop and the language of a white middle class teenager in New York City. *Journal of Sociolinguistics* 3(4). 428–442.
Drummond, Rob
2011 Glottal variation in /t/ in non-native English speech: Patterns of acquisition. *English World-Wide* 32(3). 280–308.
Drummond, Rob
2012 Aspects of identity in a second language: ING variation in the speech of Polish migrants living in Manchester, UK. *Language Variation and Change* 24(1). 107–133.
Drummond, Rob
2013 The Manchester Polish STRUT: Dialect acquisition in a second language. *Journal of English Linguistics* 41(1). 65–93.
Dubois, Sylvie & Barbara Horvath
1998 Let's tink about dat: Interdental fricatives in Cajun English. *Language Variation and Change* 10(3). 245–61.
Dubois Sylvie & Barbara Horvath
2000 When the music changes, you change too: Gender and language change in Cajun English. *Language Variation and Change* 11(3). 287–313.
Eckert, Penelope
2008a Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 453–476.
Eckert, Penelope
2008b Where do ethnolects stop? *International Journal of Bilingualism* 12(1–2). 25–42.
Fader, Ayala
2009 *Mitzvah girls*. Princeton: Princeton University Press.
Fasold, Ralph
1972 *Tense marking in Black English: A linguistic and social analysis.* Arlington: Center for Applied Linguistics.
Fought, Carmen
2003 *Chicano English in context.* New York: Palgrave Macmillan.
Friesner, Michael & Aaron Dinkin
2006 The acquisition of native and local phonology by Russian immigrants in Philadelphia. *Penn Working Papers in Linguistics* 12(2). 91–104.
Gordon, Matthew
2000 Phonological correlates of ethnic identity: Evidence of divergence? *American Speech* 75(2). 115–136.
Guy, Gregory
2013 The cognitive coherence of sociolects: How do speakers handle multiple sociolinguistic variables? *Journal of Pragmatics* 52(1). 63–71.
Hall-Lew, Lauren
2009 *Ethnicity and phonetic variation in a San Francisco Neighborhood.* Stanford, CA: Stanford University dissertation.

Harris, Roxy
2006 *New ethnicities and language use.* Basingstoke: Palgrave Macmillan.
Hinskens, Frans
2011 Emerging Moroccan and Turkish varieties of Dutch: Ethnolects or ethnic styles? In Friederike Kern & Margret Selting (eds.), *Ethnic styles of speaking in European metropolitan areas*, 103–131. Philadelphia: John Benjamins.
Hoffman, Michol F. & James A. Walker
2010 Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change* 22(1). 37–67.
Johnstone, Barbara & Scott F. Kiesling
2008 Indexicality and experience: Exploring the meanings of /aw/-monophthongization in Pittsburgh. *Journal of Sociolinguistics* 12(1). 5–33.
Konopka, Kenneth & Janet Pierrehumbert
2008 Vowels in contact: Mexican heritage English in Chicago. *Texas Linguistics Forum* 52. 94–103.
Kotsinas, Ulla-Britt
1988 Immigrant children's Swedish: A new variety? *Journal of Multilingual and Multicultural Development* 9(1–2). 129–140.
Labov, William
1962 *The social history of a sound change on the island of Martha's Vineyard, Massachusetts*. New York City, NY: Columbia University MA thesis.
Labov, William
1966 *The social stratification of English in New York City.* Washington, D.C.: Center for Applied Linguistics.
Labov, William, Sharon Ash & Charles Boberg
2006 *Atlas of North American English*. Berlin: Mouton de Gruyter.
Laferriere, Martha
1979 Ethnicity in phonological variation and change. *Language* 55(3). 603–617.
Lee, Hikyoung
2000 Korean Americans as speakers of English: The acquisition of general and regional features. Philadelphia, PA: University of Pennsylvania dissertation.
Levon, Erez
2006 Mosaic identity and style: Phonological variation among Reform American Jews. *Journal of Sociolinguistics* 10(2). 185–205.
Mendoza-Denton, Norma
2002 Language and identity. In J. K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 475–499. Malden, MA and Oxford, UK: Blackwell Publishers.
Mendoza-Denton, Norma
2008 *Homegirls: Language and cultural practice among Latina youth gangs.* Malden, MA: Blackwell Publishing.
Nagy, Naomi & Alexei Kochetov
2013 Voice onset time across the generations: A cross-linguistic study of contact-induced change. In Peter Siemund, Ingrid Gogolin, Monika Schulz & Julia Davydova (eds.), *Multilingualism and language diversity in urban areas: Acquisition, identities, space, education*, 19–38. Amsterdam: John Benjamins.
Nagy, Naomi, Joanna Chociej & Michol Hoffman
2012 Analyzing ethnic orientation in the quantitative sociolinguistic paradigm. *Language & Communication* 35. 9–26.
Newlin-Łukowicz, Luiza
2015 *Ethnicity, L1 interference, and sound change in New York City.* New York City, NY: New York University dissertation.

Newlin-Łukowicz, Luiza
2016 Co-occurrence of sociolinguistic variables and the construction of ethnic identities. *Lingua* 172/173. 100–115.
Newman, Michael & Angela Wu
2011 "Do you sound Asian when you speak English?" Racial identification and voice in Chinese and Korean Americans' English. *American Speech* 86(2). 152–178.
Nortier, Jacomine & Margreet Dorleijn
2008 A Moroccan accent in Dutch: A sociocultural style restricted to the Moroccan community? *International Journal of Bilingualism* 12(1–2). 125–142.
Payne, Arvilla Chapin
1976 *The acquisition of the phonological system of a second dialect*. Philadelphia, PA: University of Pennsylvania dissertation.
Poplack, Shana
1978 Dialect acquisition among Puerto Rican bilinguals. *Language in Society* 7(1). 89–103.
Preston, Dennis R.
2003 Systemic accommodation. In David Britain & Jenny Cheshire (eds.), *Social dialectology: In honour of Peter Trudgill*, 39–58. Amsterdam: John Benjamins.
Purnell, Thomas C., William J. Idsardi & John Baugh
1999 Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18(1). 10–30.
Quist, Pia
2008 Sociolinguistic approaches to multiethnolect: Language variety and stylistic practice. *International Journal of Bilingualism* 12(1–2). 43–61.
Rampton, Ben
1995 *Crossing: Language and ethnicity among adolescents*. London: Longman.
Rickford, John R.
1985 Ethnicity as a sociolinguistic boundary. *American Speech* 60(2). 99–125.
Rickford, John R. & Mackenzie Price
2013 Girlz II women: Age-grading, language change and stylistic variation. *Journal of Sociolinguistics* 17(2). 143–179.
Roberts, Julie & William Labov
1995 Learning to talk Philadelphian: Acquisition of short *a* by preschool children. *Language Variation and Change* 7(1). 101–112.
Roediger, David R.
2005 *Working toward whiteness: How America's immigrants became white*. Cambridge, MA: Basic Books.
Rose, Mary
2006 *Language, place, and identity in later life*. Stanford, CA: Stanford University dissertation.
Santa Ana, Otto
1991 *Phonetic simplification processes in the English of the barrio: A cross-generational sociolinguistic study of the Chicanos of Los Angeles.* Philadelphia, PA: University of Pennsylvania dissertation.
Santa Ana, Otto
1993 Chicano English and the Chicano Language Setting. *Hispanic Journal of Behavior Sciences* 15(1). 3–35.
Schleef, Erik, Miriam Meyerhoff & Lynn Clark
2011 Teenagers' acquisition of variation: A comparison of locally-born and migrant teens' realisation of English (ing) in Edinburgh and London. *English World-Wide* 32(2). 206–236.

Sharma, Devyani
2014 Transnational flows, language variation, and ideology. In Marianne Hundt & Devyani Sharma (eds.), *English in the Indian diaspora*, 215–242. Amsterdam: Benjamins.
Sharma, Devyani & Lavanya Sankaran
2011 Cognitive and social forces in dialect shift: Gradual change in London Asian speech. *Language Variation and Change* 23(3). 399–428.
Trechter, Sara & Mary Bucholtz
2001 White noise: Bringing language into whiteness studies. *Journal of Linguistic Anthropology* 11(1). 3–21.
Wolfram, Walt
1969 *A sociolinguistic description of Detroit Negro Speech*. Washington, DC: Center for Applied Linguistics.
Wolfram, Walt
2007 Sociolinguistic folklore in the study of African American English. *Language and Linguistics Compass* 1(4). 292–313.
Wolfram, Walt, Phillip Carter & Beckie Moriello
2004 Emerging Hispanic English: New dialect formation in the American South. *Journal of Sociolinguistics* 8(3). 339–358.
Wong, Amy Wing-mei
2007 Two vernacular features in the English of four American-born Chinese. *University of Pennsylvania Working Papers in Linguistics* 13(2). 217–230.
Zentella, Ana Celia
1990 Returned migration, language, and identity: Puerto Rican bilinguals in dos worlds/two mundos. *International Journal of the Sociology of Language* 84. 81–100.

Luiza Newlin-Łukowicz, Virginia (USA)

25. Language ideology

1. Language ideology as a field of inquiry
2. Linguistic boundaries: the integrity and state of the language
3. Social processes: linguistic hierarchies and regimes of language
4. Linguistic authority: authenticity and anonymity
5. Conclusion
6. References

1. Language ideology as a field of inquiry

The field of language ideology, which explores the different interactions and facets of language beliefs and use in cultural contexts, developed into a significant scholarly tradition in the late twentieth century. The field is now well-established, and its theoretical framework has facilitated a wide range of studies across multiple disciplines and geographic contexts over the past 30 years. This can be traced back to linguistic anthropological work on language ideology that began to gain momentum in the late 1970s, in particular with Silverstein's (1979: 193) contribution, in which he defines language ideol-

https://doi.org/10.1515/9783110435351-025

ogy as "sets of beliefs about language articulated by users as a rationalization or justification of perceived language structure and use". This emphasis on rationalization is based on a critical value of language ideology that is linked to the social-cognitive processes of distortion. Other scholars have conceptualized language ideology in a more neutral and all-encompassing manner, such as Rumsey (1990: 346), who defines it as "shared bodies of common sense notions about the nature of language in the world". As Woolard and Schieffelin (1994: 58) point out, these different approaches sometimes stem from divergent interpretations and theorizations of ideology itself, but they share the view that ideology is bound up with a particular social position. Irvine's (1989: 255) widely cited definition of language ideology underscores the importance of positionality: "the cultural system of ideas about social and linguistic relationships, together with their loading of moral and political interests".

Part of this research were increasingly critical views of sociolinguistics, including work that interrogated the role of linguists in the construction of linguistic reality. The essays in Joseph and Taylor (1990) consider the ways that ideological processes impact the history of linguistics, as well as how linguists contribute to the reification of named languages and to the construction of the ideal of standard language. The chapters in Blommaert (1999) extend this critique from linguistics to language in the broader sociopolitical sphere by examining the role of social actors, and especially of the more active 'ideology brokers' in language matters. This volume put the debate at the heart of sociolinguistic inquiry, and the notion of language ideological debates was formative to many ensuing studies (see Sections 2 and 3). This body of work underscores the need to foreground the analysis of discourse and stresses that there is no such thing as a neutral position. Moreover, it is not taken for granted that languages and societies are natural and clearly bounded entities. In their related discussion of research on metalanguage, Coupland and Jaworski (2004: 38–40) stress that developments in language ideology have played a key role in shaping the "end of innocence" that marks sociolinguistic research in the early twenty-first century.

At the turn of the twenty-first century, three seminal volumes on language ideology appeared as a result of multiple and intersecting exchanges at academic events focused on this theme. These volumes moved the field forward on a conceptual level, and crucially, made work on language ideology more widely known beyond the ambit of anthropological linguistics. Comprised of chapters dealing with a wide range of sites, in the first of these volumes (Schieffelin, Woolard, and Kroskrity 1998), Woolard (1998: 3) highlights the role of language ideology as a "mediating link between social forms and forms of talk". The second (Kroskrity 2000b) explores regimes of language, and in his introduction, Kroskrity (2000a) argues for treating language ideology as a cluster concept in order to explore its multifaceted nature. Additionally, this volume contains a chapter by Irvine and Gal (2000) on linguistic differentiation, which develops an influential conceptual model based on the semiotic processes of iconization, fractal recursivity, and erasure. The third volume (Gal and Woolard 2001b) considers the interface between language and publics, as well as the notion of linguistic authority, discussed in Section 4. Taken as a whole, research on language ideology emphasizes the social and personal dimensions of how language ideologies shape, and are shaped by, linguistic and social structures through the means of agency, social positioning, and partiality, as well as rejection and contestability. Linguistic anthropologists were not alone in considering ideological aspects of language, in particular the construction of boundaries, hierarchies, and publics,

but this core body of work referred to as 'language ideologies' would go on to have a transformative impact on the field of sociolinguistics in the early part of the twenty-first century.

In summary, language ideology is a complex concept and analytical tool. Language ideologies are not unilateral, and scholars offer a range of definitions and interpretations, each with their own nuances. While many of these only entail minor contestations, e.g. whether the primary focus of the field should be on linguistic structures (linguistic ideology) or on language and social order (language and ideology), there is a consensus about the core concept − most noticeably that language ideologies do not pertain exclusively to language matters. Typically, questions of language ideology are bound up with questions of power, inequality, and struggle. Ideology is best understood as an active process that is subject to interpretation and negotiation through discourse, and related to socialization and cognition. Ideologies rely on underlying belief systems shared by group members to varying degrees, which in turn are mobilized by certain individuals, and legitimated and reproduced within institutional frameworks. Examples of language ideologies include: what linguistic forms are deemed (in)correct; what or how individuals and collectives should speak; the arbitrary value and legitimacy allocated to some language varieties over others; and how identities are shaped, constructed, and performed on the basis of the language(s) one speaks or acquires. The following sections explore ideologies concerning linguistic boundaries and those ideologies pertaining to language and broader social processes.

2. Linguistic boundaries: the integrity and state of the language

The notion that languages exist as clearly identifiable and bounded objects constitutes a widespread and fundamental language ideological belief. To be sure, the construction of named languages functions similarly to that of other categories, such as ethnicities and nations, all of which can be interpreted and instrumentalized in various ways due to their potential malleability and their situatedness in given social and political contexts. The demarcation of named languages is a process which has been closely bound up with European nationalisms which fuse nation and state, became prominent during the long nineteenth century, and spread on a global scale via colonization. Irvine and Gal (2000) underscore the ways that semiotic processes inform the construction of ideological representations of linguistic difference and linguistic categorization. In their critical discussion of linguistic classifications of the Nguni languages of southern Africa, the languages of Senegal, and the linguistic varieties of Macedonia, Irvine and Gal (2000: 76) stress the need for linguistic research, in particular that on language contact, to "shift attention to linguistic differentiation rather than community" and, with reference to Anderson ([1983] 1991), they underscore that "homogeneous language is as much imagined as is community".

Language ideological research has had significant impact on problematizing the notion of homogeneous language and the reification of named languages. It has explored the construction of standardized national languages, as well as the categorization of languages and dialects, in relation to the deeply entrenched belief that nation-states require their own language to justify their autonomy and to construct authority within the polity.

This line of inquiry bears affinities to Bourdieu's (1991: 43–65) critique of linguistics and his discussion of the legitimate language and symbolic power. In a related vein, the formative monograph by Milroy and Milroy ([1985, 1991] 1999) foregrounds the ways that the standard language, while intangible and characterized as an ideological construct, is a powerful ideal that plays a key role in the reproduction of social inequality. The standard language ideology is the idea that there exists a fixed, unchangeable, and homogeneous variety of a language, which also tends to be regarded as the most highly valued variety. This in turn underpins prescriptivist views on how the language ought to be, including accent and grammar, giving rise to notions of 'good' and 'bad' language based on institutionalized practices concerning what is deemed 'correct' and 'incorrect'. Cameron (2004: 313) points out that "metalinguistic resources seem very often to be deployed to connect various aspects of linguistic behaviour to a larger moral order". She asserts that linguists also bear responsibility in the propagation of language ideologies (Cameron 2004: 315–317), a point also highlighted by Milroy (2001).

Linguistic purism constitutes a further means of regulating the moral order, targeting items that are perceived to be foreign and intrusive in the linguistic system. Linguistic purism manifests itself in myriad socio-political contexts, and there is evidence of this phenomenon in historical and contemporary timeframes alike (Jernudd and Shapiro 1989; Langer and Davies 2005; Thomas 1991). Issues of purism become more ideologically complex in the context of language revitalization and promotion of minoritized languages. One issue explored by Jaffe (1999) in her early work on language ideologies on Corsica is the way in which the Corsican language is strategically constructed as distinct from Italian. Jaffe notes how activists engage in linguistic purism, targeting Italian lexical items, so that Corsican can be perceived as a language in its own right and stand as a valid competitor to the hegemony of French. In a recent study concerning the use and maintenance of vernacular Maya on the Yucatan peninsula, Cru (2016: 127) discusses how Maya speakers experience ideological tensions due to the fact that "some degree of purism may be a positive component for legitimation, [while] extreme purist attitudes may have the detrimental effect of paving the way for language shift". He finds that certain speakers are adopting a somewhat more flexible approach. Similarly, Jaffe (2005) describes how the more recent polynomic approach to Corsican, which is seemingly tolerant of linguistic variation, is implemented in schools and allows a flexible use of Corsican, although this is mostly limited to a small set of spellings. At the same time, there exists the belief that Corsican and French need to be kept separate. Ideologies of purism, standardization, and orthography bear substantial weight with regard to issues of language and education, and as a result, language ideological work has proliferated in this area of research.

At times, orthography itself becomes the target of language ideological debates, and the controversy surrounding the German orthographic reform of 1996 is an interesting case in point. Although this reform of spelling and punctuation was relatively minor in nature, for example, by restricting the *Eszett* (ß) to appear only after long vowels and diphthongs, it triggered significant debate and opposition that culminated in a hearing in the German Federal Constitutional Court. Drawing on Blommaert (1999), Johnson (2005) provides an in-depth analysis of the positions of various 'ideology brokers' in this debate, i.e. the complainants, judges, and linguists, and explores how language ideologies are justified and contested, while also considering how the debate relates to the longer historical trajectory and construction of nationhood. In a subsequent volume fo-

cused on orthography and script (Jaffe et al. 2012), Sebba (2012: 15) underscores the importance of interdisciplinarity, stating that "script and spelling are often just the outward faces of deep-seated questions of ideology and identity". This statement resonates directly with a fundamental point concerning research on language ideologies, in that this body of work stresses how beliefs about language are never about language alone, even if on the surface the focus appears to be on specific linguistic forms or on the construction of linguistic boundaries.

3. Social processes: linguistic hierarchies and regimes of language

As cultural systems of ideas and feelings, as well as of norms and values, language ideologies inform the way that people view not only language itself, but also broader relationships between language and society. Because ideologies constitute simplified, or even distorted, representations of reality, they tend to shape stereotypical thinking and can even impact on policies that influence social inclusion and exclusion. According to Kroskrity (2000a), language ideologies are characterized by four interrelated dimensions: (1) they represent perceptions of language and discourse that are constructed in the interests of a specific group; (2) they are multiple, rather than fixed or unitary; (3) people usually display varying degrees of awareness of these language ideologies; and (4) language ideologies mediate between social structures and forms of talk. Language ideologies are complex, and often overlapping and interconnected. Related notions of language hierarchies and regimes are based on the belief that language can be labelled and divided into languages, dialects, and so on, which are then subsumed into a hierarchy, with languages often being looked upon as superior to dialects, and additionally, certain languages being given a higher status as the national (or official) language of a polity or community. Bound up with potential normative power and imbued with vested interests, language ideologies shape the dynamics of group membership and the negotiation of social boundaries.

Language serves to naturalize group boundaries, and this process underpins the ideal of a linguistically and culturally homogeneous society, which Blommaert and Verschueren (1998: 194–195) refer to as the "dogma of homogeneism". This homogeneity not only pertains to the way people speak a language, but also underpins the idea that everyone should speak a given language in a particular society. Speaking a particular language or variety positions individuals in relation to groups, as well as within the broader sociopolitical sphere. Moreover, language is an integral part of how that world itself is structured and regarded. Gal and Woolard (2001a: 1) note how language ideologies and "the work of linguistic representation produces not only individualised speakers and hearers as the agents of communication, but also larger, imagined social groupings". One of the most widespread language ideologies is 'one nation, one language', which directly equates a language with a nation, on the basis that the linkage between these concepts is natural, timeless, and indisputable. Language is sometimes equated with territory, and the link between language and national identity is essentialized to the point where other forms of language use are rendered invisible. Similar to the notion of bounded languages, the idealized congruence between one nation and one language became widespread in Europe in the long nineteenth century and proliferated globally in connection with colo-

nization, and notably, all the more so in the post(−)colonial period. In tandem with nation-state congruence, language policies have frequently led to the marginalization of linguistic minorities, sometimes by pushing them to the periphery of the polity, and other times by eradicating the use of the minority language altogether.

Even in countries that officially recognize more than one language, such as Singapore, Bokhorst-Heng (1999: 238) notes that policies on multilingualism are based on "a curious twist: 'mother-tongue' is defined according to one's father's ethnicity", multilingualism is managed in a prescribed and compartmentalized manner, and the nation is constructed as a "patchwork of internally homogenized communities" (Bokhorst-Heng 1999: 262). Significantly, this homogenization also involves further layers of language ideological work, including linguistic purism and the misrecognition of Chinese dialects, fueled by the Speak Mandarin Campaign as a means of homogenizing the Chinese community. In a related vein, Heller (1999) draws attention to and problematizes the notion of 'two solitudes' in her discussion of language ideological debates in Canada. She describes tensions in Quebec concerning competing ethnic and civic nationalisms, and how "minority francophones seek to navigate their way between these two competing visions, neither one of which seems to really represent their interests" (Heller 1999: 167). The globalizing context has complicated minority rights, and language ideological inquiry has been one means of equipping researchers to engage with the increasing level of complexity inherent to these issues.

A further area where language ideological approaches have been influential is language and migration, in particular relating to language policies affecting migration. Since the turn of the twenty-first century, multiple polities in the global North have introduced new forms of citizenship legislation that in many cases include language and/or civics tests. Research informed by language ideologies and regimes of language has explored the motivations for and consequences of recently implemented language requirements and/or testing procedures (Extra, Spotti, and Van Avermaet 2009; Hogan-Brun, Mar-Molinero, and Stevenson 2009). This research has critically engaged with the gate-keeping dimension of these measures, as well as the related valorization and stigmatization of certain linguistic repertoires and identities, and analyzed the ways in which this recent legislation is justified by social actors at the level of individual states. Some studies have also investigated tensions inherent to these policies, as well as the related contestation of the policies. For instance, Horner (2015) discusses how a certain degree of uncertainty around the 'linguistic authority' of Luxembourgish gave rise to disputes concerning the introduction of formalized language testing for citizenship in the trilingual country of Luxembourg. Unlike many other European countries, where the issue of which language plays the role of the language of integration is considered 'common sense' knowledge, considerable efforts have been made to establish Luxembourgish in this function, and related tensions give rise to language ideological debates in the public sphere.

4. Linguistic authority: authenticity and anonymity

Language ideological research throws light on how some language varieties are seen to be prestigious, useful, and legitimate, while others are perceived to be insignificant, unimportant, and even useless. In their work on languages and publics, Gal and Woolard

(2001a: 2–5) expand upon research on language ideologies by discussing how representations of linguistic phenomena come to gain credibility and become influential, and also how linguistic ideologies and practices are pivotal in the formation of political authority. They explain that linguistic authority is established from two different perspectives: (1) emphasizing the "socially locatable" nature of language, rooted in primordialism and territorial belonging; and (2) emphasizing the "aperspectival objectivity" of language, based on universality and prescriptive norms. The former prong resonates with the 'one nation, one language ideology', and the latter with the standard language ideology as discussed above. Gal and Woolard (2001a) stress that either perspective may be invoked to construct linguistic authority, and that the relationship between the two perspectives is often complex. They also underscore that the notion of "aperspectival objectivity", or the "view from nowhere", is pivotal in the process of constructing languages and publics alike (Gal and Woolard 2001a: 4). In this vein, they elucidate on two interrelated strategies that inform research on language ideologies: "[t]he first strategy focuses on the different social sources of authority for diverse definitions of language phenomena. The second strategy analyzes the ways in which particular linguistic practices and beliefs about languages buttress the legitimacy of specific political arrangements" (Gal and Woolard 2001a: 10).

Extending this work on linguistic authority, Woolard (2008) applies the aforementioned two-pronged means of constructing authority directly to named languages – more specifically, to language politics in Catalonia – mapping the former 'socially locatable' aspect onto authenticity and the latter 'aperspectival objectivity' onto anonymity. Woolard (2008) shows how there has been a shift from the construction of Catalan as a national symbol (ideology of authenticity) to that of Catalan as everybody's language and yet nobody's language in particular (ideology of anonymity), while also highlighting the tensions inherent to this process with regard to regional and/or small languages. Language ideological approaches enable researchers to identify such tensions and to understand shifts in language policies of states, regions, and so on.

Ideologies of linguistic authenticity mainly concern the value of a particular language variety and its relationship to a community or speaker. In authenticity-based discourses, origins and beginnings define essence. Thus, a linguistic variety must possess locality and be rooted in a specific social or geographical space to be perceived as linguistically authentic. In other words, it must necessarily be 'from somewhere'. Similarly, a linguistic variety is inauthentic when origin or locality cannot be traced (Woolard 2016). Ideologies of linguistic authenticity are evident, for example, in many discourses surrounding second language learning. These assume that 'new speakers' (O'Rourke, Pujolar, and Ramallo 2015) and second-language learners are not legitimate speakers, and their methods of learning (i.e. classroom language learning) are not authentic or 'natural'. A study by King and Hermes (2014) illustrates how passive methodologies of language learning are favored due to speakers' fear of practicing and speaking the indigenous language inauthentically in conversation. Authentic language is thus the natural, the first, and the original, whereas inauthentic language is the artificial, the second, and the learned.

In juxtaposition, for ideologies of linguistic anonymity, a language is a "neutral vehicle of communication" (Woolard 2016: 7) which belongs to nobody or to no one in particular, and thus is equally available to all. Unlike in ideologies of linguistic authenticity, social origins are often erased (Irvine and Gal 2000); if not, they are represented as transcended (Woolard 2016). In contrast to minoritized languages, hegemonic languages

in modern western societies often rest their authority on a conception of anonymity and a voice from 'nowhere'. This is indeed the case for Spanish, which is framed as a globalized language and built on an ideology of anonymity. Spanish is increasingly seen as a commodity (Mar-Molinero 2006), and marketed as a means of access to a global and cosmopolitan community, with important economic and social benefits. Moreover, many sociolinguistic case studies have shown how ideologies of linguistic anonymity may reinforce already dominant languages and consolidate their authority. As a consequence, this dominance often becomes invisible, and in this way, the superordinate position is naturalized and unquestioned.

The relationship between anonymity and authenticity is not dichotomous, but interrelated; they are defined in relation to each other as poles on an axis of linguistic differentiation (Gal 2012). In her recent work on ideologies of linguistic authority, Woolard (2016) has extended the analysis of this relationship. She suggests that the two share a common basis in an ideology of sociolinguistic naturalism, which she defines as "the assumption that a linguistic form exists independent of willful human intervention and that it naturally and directly corresponds to a social state of affairs" (Woolard 2016: 7). This sociolinguistic naturalism (and by extension, the ideologies of anonymity and authenticity) rejects the agency of speakers and their active role in constructing the social and linguistic world around them. As discussed in Section 1, the field of sociolinguistics and linguistic anthropology has for a long time now been challenging the traditionalist and essentialist view of linguistic naturalism, as can be seen in many recent works emphasizing personal agency and resistance. Emerging alternative interpretations of ideologies of authenticity are one such example of this trend. Indeed, conceptualizations of authenticity are not always, or necessarily, founded in (sociolinguistic) naturalism. Woolard (2016) points out that it can be a self-aware or reflexive stance based on personal agency and choice.

This non-naturalist understanding of authenticity, or "project authenticity" (Woolard 2016: 38), has become significant for analyzing some modern European contexts, such as Catalonia. In twenty-first century Catalonia, Woolard (2016: 300) finds "an escape from the binary logic of authenticity versus anonymity altogether into forms of rooted cosmopolitanism, and toward a post-natural, goal-oriented sense of authenticity as project rather than origin". Upon re-interviewing participants from a previous study (Woolard 1987) twenty years later, Woolard finds that most of them adopt a personalistic attitude towards language based on freedom of choice. Naturalist understandings of authenticity, however, are still particularly prominent among immigrant students. Recent research shows how students from Latin American backgrounds perceive Catalan as the 'exclusive preserve of the Catalans' (Patiño-Santos 2018: 82), and the Catalans as owners of their language and its authenticity. As a result of these language ideologies, many immigrant students refuse to speak Catalan, and see more utility and mobility in other named languages, such as Spanish and English (Newman, Patiño-Santos, and Trenchs-Parera 2013).

5. Conclusion

Since the turn of the twenty-first century, research on language ideology has had a transformative impact in the fields of linguistic anthropology and sociolinguistics. It has also been influential on research on language and education, language and migration,

and language in the media. In tandem with constructivist impulses in the social sciences, critiques of linguistics, and the discursive turn in sociolinguistics and language policy, this body of scholarship has led researchers to explore the construction of boundaries and linguistic differentiation, rather than taking bounded languages and communities as the focus of research. This development is crucial and timely in the globalizing context, which is characterized by augmented forms of mobility and the rapid diversification of sociolinguistic practices (see Blommaert 2010; Heller 2011). Moreover, language ideological research has brought the notions of debates and publics to the forefront, which underscores the need for studies on language issues to engage with interdisciplinary work on structure and agency, as well as mediality and multimodality. As Lim and Ansaldo (2016: 2) point out, language contact is "the study of speakers in contact", which includes the role that language plays in the negotiation of identity and power relationships, as well as how language ideologies underpin these social processes. In this way, language ideological approaches highlight three key points that would benefit future work in contact linguistics: language is inherently social and political, debates about language are never about language alone, and, crucially, a neutral point of view is an illusion.

6. References

Anderson, Benedict
1991 [1983] *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.

Blommaert, Jan (ed.)
1999 *Language ideological debates*. Berlin & New York: Mouton de Gruyter.

Blommaert, Jan
2010 *The sociolinguistics of globalization*. Cambridge: Cambridge University Press.

Blommaert, Jan & Jef Verschueren
1998 *Debating diversity: Analysing the discourse of tolerance*. London: Routledge.

Bokhorst-Heng, Wendy
1999 Singapore's Speak Mandarin campaign: Language ideological debates in the imagining of the nation. In Jan Blommaert (ed.), *Language ideological debates*, 235–265. Berlin & New York: Mouton de Gruyter.

Bourdieu, Pierre
1991 *Language and symbolic power*. Harvard & Cambridge: Harvard University Press.

Cameron, Deborah
2004 Out of the bottle: The social life of metalanguage. In Adam Jaworski, Nikolas Coupland & Dariusz Galansinski (eds.), *Metalanguage: Social and ideological perspectives*, 311–321. Berlin and New York: Mouton de Gruyter.

Coupland, Nikolas & Adam Jaworski
2004 Sociolinguistic perspectives on metalanguage: Reflexivity, evaluation and ideology. In Adam Jaworski, Nikolas Coupland & Dariusz Galansinski (eds.), *Metalanguage: Social and ideological perspectives*, 15–51. Berlin & New York: Mouton de Gruyter.

Cru, Josep
2016 Shifting language ideologies among young Maya professionals: Overcoming purism in Yucatan. *Critical Multilingualism Studies* 4(2). 111–132.

Extra, Guus, Massimiliano Spotti & Piet Van Avermaet (eds.)
2009 *Language testing, migration and citizenship: Cross-national perspectives on language regimes*. London: Continuum.

Gal, Susan
2012 Sociolinguistic regimes and the management of "diversity". In Alexandre Duchêne & Monica Heller (eds.), *Language in late capitalism: Pride and profit*, 22–42. New York: Routledge.
Gal, Susan & Kathryn A. Woolard
2001a Constructing languages and publics: Authority and representation. In Susan Gal & Kathryn A. Woolard (eds.), *Languages and publics: The making of authority*, 1–12. Manchester & Northampton: Routledge.
Gal, Susan & Kathryn A. Woolard (eds.)
2001b *Languages and publics: The making of authority*. Manchester & Northampton: Routledge.
Heller, Monica
1999 Heated language in a cold climate. In Jan Blommaert (ed.), *Language ideological debates*, 143–170. Berlin & New York: Mouton de Gruyter.
Heller, Monica
2011 *Paths to post-nationalism: A critical ethnography of language and identity*. Oxford: Oxford University Press.
Hogan-Brun, Gabrielle, Clare Mar-Molinero & Patrick Stevenson (eds.)
2009 *Discourses on language and integration: Critical perspectives on language testing regimes in Europe*. Amsterdam & Philadelphia: John Benjamins Publishing Company.
Horner, Kristine
2015 Language regimes and acts of citizenship in multilingual Luxembourg. *Journal of Language and Politics* 14(3). 359–381.
Irvine, Judith T.
1989 When talk isn't cheap: Language and political economy. *American Ethnologist* 16(2). 248–267.
Irvine, Judith T. & Susan Gal
2000 Language ideology and linguistic differentiation. In Paul V. Kroskrity (ed.), *Regimes of language: Ideologies, polities and identities*, 35–83. Santa Fe & Oxford: School of American Research Press.
Jaffe, Alexandra
1999 *Ideologies in action: Language politics on Corsica*. Berlin & New York: Mouton de Gruyter.
Jaffe, Alexandra
2005 La polynomie dans une école bilingue corse: Bilan et défis [Polynomy in a Corsican bilingual school: Results and challenges]. *Marges linguistiques* 10. 282–300.
Jaffe, Alexandra, Jannis Androutsopoulos, Mark Sebba & Sally Johnson (eds.)
2012 *Orthography as social action: Scripts, spelling, identity and power*. Boston & Berlin: Mouton de Gruyter.
Jernudd, Björn H. & Michael J. Shapiro (eds.)
1989 *The politics of language purism*. Berlin & New York: Mouton de Gruyter.
Johnson, Sally
2005 *Spelling trouble? Language, ideology and the reform of German orthography*. Clevedon: Multilingual Matters.
Joseph, John E. & Taylor J. Talbot
1990 *Ideologies of language*. London & New York: Routledge.
King, Kendall A. & Mary Hermes
2014 Why is this so hard? Ideologies of endangerment, passive language learning approaches, and Ojibwe in the United States. *Journal of Language, Identity and Education* 13(4). 268–282.

Kroskrity, Paul V.
2000a Regimenting languages: Language ideological perspectives. In Paul V. Kroskrity (ed.), *Regimes of language: Ideologies, polities and identities*, 1–34. Santa Fe & Oxford: School of American Research Press.
Kroskrity, Paul V. (ed.)
2000b *Regimes of language: Ideologies, polities and identities*. Santa Fe & Oxford: School of American Research Press.
Langer, Nils & Winifred V. Davies (eds.)
2005 *Linguistic purism in the Germanic languages*. Berlin & New York: Mouton de Gruyter.
Lim, Lisa & Umberto Ansaldo
2016 *Languages in contact*. Cambridge: Cambridge University Press.
Mar-Molinero, Clare
2006 The European linguistic legacy in a global era: Linguistic imperialism, Spanish and the Instituto Cervantes. In Clare Mar-Molinero & Patrick Stevenson (eds.), *Language ideologies, policies and practices: Language and the future of Europe*, 76–91. Hampshire & New York: Palgrave Macmillan.
Milroy, James
2001 Language ideologies and the consequences of standardization. *Journal of Sociolinguistics* 5(4). 530–555.
Milroy, James & Lesley Milroy
1999 [1985, 1991] *Authority in language: Investigating Standard English*. London & New York: Routledge.
Newman, Michael, Adriana Patiño-Santos & Mireia Trenchs-Parera
2013 Linguistic integration of Latin American immigrants in Catalonia and their responses to the implementation of educational language policies. *International Journal of Bilingual Education and Bilingualism* 16(2). 195–209.
O'Rourke, Bernadette, Joan Pujolar & Fernando Ramallo
2015 New speakers of minority languages: The challenging opportunity – Foreword. *International Journal of the Sociology of Language* 231. 1–20.
Patiño-Santos, Adriana
2018 "No-one told me it would all be in Catalan!" – Narratives and language ideologies in the Latin American community at school. *International Journal of the Sociology of Language* 250. 59–86.
Rumsey, Alan
1990 Wording, meaning and linguistic ideology. *American Anthropologist* 92(2). 346–361.
Schieffelin, Bambi B., Kathryn A. Woolard & Paul V. Kroskrity (eds.)
1998 *Language ideologies: Practice and theory.* Oxford & New York: Oxford University Press.
Sebba, Mark
2012 Orthography as social action: Scripts, spelling, identity and power. In Alexandra Jaffe, Jannis Androutsopoulos, Mark Sebba & Sally Johnson (eds.), *Orthography as social action: Scripts, spelling, identity and power*, 1–19. Boston & Berlin: Mouton de Gruyter.
Silverstein, Michael
1979 Language structure and linguistic ideology. In Paul R. Clyne, William F. Hanks & Carol L. Hofbauer (eds.), *The elements: A parasession on linguistic units and levels*, 193–247. Chicago: Chicago Linguistic Society.
Thomas, George
1991 *Linguistic purism*. London: Longman.
Woolard, Kathryn A.
1987 Codeswitching and comedy in Catalonia. *Pragmatics* 1(1). 106–122.

Woolard, Kathryn A.
1998 Introduction: Language ideology as a field of inquiry. In Bambi B. Schieffelin, Kathryn A. Woolard & Paul V. Kroskrity (eds.), *Language ideologies: Practice and theory*, 3–47. Oxford & New York: Oxford University Press.
Woolard, Kathryn A.
2008 Language and identity choice in Catalonia: The interplay of contrasting ideologies of linguistic authority. In Kirsten Suselbeck, Ulrike Mühlschlegel & Peter Masson (eds.), *Lengua, nación e identidad: La regulación del plurilingüismo en España y América Latina* [*Language, nation, and identity: The regulation of multilingualism in Spain and Latin America*], 302–324. Berlin: Bibliotheca Ibero-Americana.
Woolard, Kathryn A.
2016 *Singular and plural: Ideologies of linguistic authority in 21st century Catalonia*. Oxford: Oxford University Press.
Woolard, Kathryn A. & Bambi B. Schieffelin
1994 Language ideology. *Annual Review of Anthropology* 23. 55–82.

Kristine Horner, Sheffield (UK)
Andrew F. Bradley, Sheffield (UK)

26. Speech community

1. Introduction
2. Linguistic, language, and speech communities
3. Shared norms and ideologies
4. Language variation and contact
5. Community and place
6. Speech
7. Beyond speech communities: communities of practice
8. Conclusions
9. References

1. Introduction

The notion of speech community is central to the fields of sociolinguistics, variationist (socio)linguistics, sociology of language, linguistic anthropology, and contact linguistics. Hymes (1964: 385) described it as a "fundamental concept for the relation between language, speech, and social structure", which, according to Wardhaugh (2006: 119), "has proved to be invaluable in sociolinguistic work". However, definitions of the concept vary, and Wardhaugh (2006: 119) observes that there is "a certain 'fuzziness' as to its precise characteristics". As noted by Austin and Sallabank (2014: 8), the concept is deeply problematic: what are its boundaries, what is its scale, and is it still relevant to current research paradigms?

This chapter reviews approaches to speech community, its relevance to contact linguistics, and the ways in which the concepts of community, speech, and speech community are treated in more recent literature, including the related concept 'communities of practice'. It attempts to address both traditional and newer ways of approaching the topic in a constructively critical fashion.

https://doi.org/10.1515/9783110435351-026

2. Linguistic, language, and speech communities

This chapter does not attempt a full history of the term 'speech community', which has already been provided admirably by several authors, notably Patrick (2003), Morgan (2005, 2014), Mendoza-Denton (2011), and Madera (1996). Nevertheless, some historical context is necessary for discussion of the concept and its implications.

As far as is traceable, the term 'speech community' was first used in English by Bloomfield (1933: 29), who defined it as "a group of people who use the same set of speech signals". In this tradition, which can be seen as following German conventions (Ammon 1995: 32), a speech community is seen as relatively homogenous and defined according to linguistic criteria:

> *Eine Sprachgemeinschaft umfasst alle Individuen gleicher Muttersprache, genauer: die Muttersprachler aller Varietäten ein und derselben Sprache.*
> [A speech community encompasses all individuals with the same mother tongue, or more precisely: the native speakers of all varieties of one and the same language; translation mine]

Weisgerber (1967: 30) traces this usage back to 1840, which tallies with Ammon's (1995: 18–19) explanation that German and Italian nationalists utilized the notion of *Sprachvolk*, or linguistic ethnicity, in support of a single nation-state. This definition continued in use throughout the following decades (e.g. Corder 1973; Gumperz [1968] 1971; Lyons 1970; Weinreich 1953), and continues to be used now, especially by linguists who focus on language as a system.

Sociolinguistic use of the term 'speech community' developed in tandem with Labov's ([1966] 2006, 1972) studies of speech patterns in urban communities, so it is closely linked to Labovian variationist sociolinguistics, especially with regard to its focus on social norms relating the use of different styles, registers, and/or accents to social stratification. Variationism has been criticized in recent years for pre-ordaining social categories. According to Eckert (2006: 110), for example, "[t]he speech community perspective views heterogeneity as based in a geographically defined population, and structured by broad and fundamental social categories, particularly class, gender, age, race and ethnicity". Although sociolinguistics, like the concept of speech community, has its roots in studies observing systematic links between linguistic variation and social stratification, the field has developed beyond its origins. Hymes (1968) included speech community in his ethnography of communication; there is no inherent contradiction between the notion of speech community and ethnographic research, where social categories and ways of speaking are not pre-defined. Labov's early work happened to investigate systematic variation in speech communities often thought of as monolingual (despite the over 800 languages present in New York; Endangered Language Alliance 2012), but it was recognized early on that "there are no a priori grounds which force us to define speech communities so that all members speak the same language" (Gumperz 1962: 30). This chapter argues that some of the criticisms of the concept of speech community can be seen as 'straw men', stressing alleged deficits to favor another view.

The issue of identifying and defining 'a language' as a research object cannot be disengaged from the discussion of (speech) community. Moreover, the concepts of language, community, and speech community have been constructed, deconstructed, and debated intensely and in similar ways. Pennycook (2010) questions assumptions about languages as systems or as countable entities and suggests instead that language emerges

from the activities it performs. He views language as an activity rather than a structure, as a material part of social and cultural life rather than an abstract entity. As Austin and Sallabank (2014: 14) and Whaley (2011) note, there are frequently conflicting beliefs and ideologies within speech (and language) communities regarding the definition of language, its status, domains, functions, and policy, and who has the authority or legitimacy to decide any of these. For a more thorough discussion of the notion of language boundaries, see Chapter 23, this volume.

Gumperz (1962: 31) introduced the term *linguistic community*, which he defined as "a social group which may be either monolingual or multilingual, held together by frequency of social interaction patterns and set off from the surrounding areas by weaknesses in the lines of communication". This corresponds to the social construct-based definition of speech community most widely accepted in sociolinguistics and linguistic anthropology (Hymes 1974; Labov 1972). However, no community uses only one language variety: all incorporate variation in usage according to age, register, etc. According to Morgan (2005: 8), Hymes considered the question of boundaries essential in order to recognize that communities are not by definition fixed units. Hymes (1974: 47) expressly distinguished between speech community and language: "Speech community […] postulates the unit of description as a social, rather than a linguistic, entity". Hymes' definition of speech community is related to his concept of communicative competence − the knowledge a speaker must have to function as a member of a social group − deduced through the ethnography of communication (Johnstone and Marcellino 2010).

3. Shared norms and ideologies

Labov (1972: 120−121) asserted that "[t]he speech community is not defined by any marked agreement in the use of *language elements*, so much as by participation in a set of shared norms". Spolsky (2004: 14) adds that this "general set of beliefs about appropriate language practices [can form] a consensual ideology, assigning values and prestige to various aspects of the language varieties used in it". This articulates the link between shared norms, or understandings of which ways of speaking and communicating are appropriate in certain contexts, and language ideologies (Schieffelin, Woolard, and Kroskrity 1998), which, according to Steger (2003: 93), are "defined as a system of widely shared ideas, patterned beliefs, guiding norms and values, and ideals accepted as truth by a particular group of people". A speech community is thus made up of people who share folk-linguistic ideologies (Nieldzielski and Preston 2003) of what language is and what it indexes: which varieties are suitable for which domains, what social hierarchies are associated with particular ways of speaking, etc. Although norms are not the same as ideologies, they reflect underlying societal beliefs.

The perceived appropriateness of particular ways of speaking for particular domains is one of the shared norms associated with speech community. The concept of domains was one of Fishman's contributions to the field, as pointed out by García and Schiffman (2006: 18). Fishman (1967) also extended the definition of diglossic speech communities (Ferguson 1959) to include unrelated languages, arguing that intra-linguistic variation can be seen as part of a continuum with societal bi/multilingualism. Fishman was aware that in societies where monolingual ideologies dominate, ways of speaking associated with more powerful domains and groups have a tendency to influence societal norms, to the detriment of multilingual practices, as epitomized in Fishman (1991).

As Austin and Sallabank (2014: 13) observe, shared norms in a speech community do not necessarily imply that members have somehow agreed on a particular viewpoint, or that the "general set of beliefs" is held consciously. Labov (1972) differentiated between overt (or dominant) and covert norms: in effect, covert norms constitute an unofficial language policy which regulates expectations of language behavior, but which will probably not be explicitly formulated (Schiffman 1996). Although they may not correspond with prestige standards (Morgan 2005: 9), these norms and beliefs nevertheless reflect and reinforce hegemonic power relations in society, through assumptions about social roles and values reflected in language practices (Bourdieu 1991).

Despite Spolsky's and Labov's characterization of speech communities as entailing consensual norms, Rickford (1993: 4) points out the inconsistency that "while emphasizing potential heterogeneity in language use, sociolinguists looked for conformity on another level, treating it as axiomatic that members of a speech community would be united with respect to shared norms for the use of language". Rickford describes a plantation community in Guyana, where matched guise tests found that speech patterns reflected competition and conflict in the community. Morgan (2014: 13), also discussing plantation contexts where creole languages emerged, notes that "[t]hese situations also provide an opportunity to illuminate the sites of contention in which creole language speakers and descendants negotiate and seek power". Morgan (2014: 11, 33) points out that Milroy and Milroy (1992: 3) argue that linguistic norms reflect social hierarchization, and so "are more readily interpretable as evidence of conflict and sharp divisions within society than as evidence of consensus". Bucholtz (1999) echoes this view (see below).

4. Language variation and contact

The ongoing expansion of speech community from a group of people "sharing a denotational code" (Silverstein 1996: 126) to a community of people with multilingual repertoires can be traced from the acceptance of systematic variation in usage and style within the boundaries of a language, to encompass style-shifting, translanguaging, and creolization in a repertoire of socially defined codes. Morgan (2005, 2014) emphasizes the importance of the study of language contact, especially creoles, in challenging early definitions of speech community predicated on monolingualism. Le Page and Tabouret-Keller (1985: 115) describe how Le Page's research found discrepancies between the Labovian model of a speech community based on "speakers of the 'same language'", and the heterogeneous contexts he was studying in Belize.

Winford (2003), following Loveday (1996), outlines a typology of language contact settings, from relatively homogenous monolingual communities (with little direct contact with other language speakers) to varying degrees of bi- and multilingualism (stable and unstable). Unstable contexts may lead to language shift or to the emergence of new vernaculars, while in more stable settings there may be high degrees of individual multilingualism. Monolingual speech communities are, however, increasingly rare in a world characterized by superdiversity and transnational flows (Pennycook 2007), where speech communities are constantly enriched by contact with new members from distant geographical locations. Reinterpreting what it means to sound local, migrants in the Norwegian city of Bergen, for example, introduce innovative norms (Kerswill 2001).

Although some documentary and descriptive linguists attempt to rewind language contact in order to distinguish an idealized, pre-contact *ancestral code* (Childs, Good, and Mitchell 2014), increasingly the linguistic ecologies, repertoire(s) and practices of highly multilingual speech communities are the subject of research. Lüpke (2016) examines multilingual practices in a village in Lower Casamance (Senegal), based on rich ethnographic description of two households whose speech is characterized by intensive code-switching. In a similar approach, Di Carlo and Good (2014) stress the need to document a multilingual speech community's full linguistic ecology in Cameroon, while Migge and Léglise (2013) observe emergent mixed languages in French Guiana. In all these cases, intensive language contact, rather than a particular language, is seen as a key factor in local identities. The use of more than one language is thus recognized as being normal *within* a multilingual speech community, rather than necessarily an intergroup phenomenon: i.e. language contact does not happen solely due to contact *between* two language communities.

5. Community and place

The term 'community' is of course not restricted to sociolinguistics and linguistic anthropology. In sociology, the concept of community has traditionally been strongly associated with *place*. Neal (2012: n.p.) pithily summarizes the development of the concept:

> First, a community is a group of people who interact with one another, for example, as friends or neighbors. Second, this interaction is typically viewed as occurring within a bounded geographic territory, such as a neighborhood or city. Third, the community's members often share common values, beliefs, or behaviors.

According to Gold (2005: 2), the term 'community' "evokes a group of people who have something in common and who are actively engaged with one another in a benign fashion", which overlaps with the concept of communities of practice (Section 7). The term 'community' has overwhelmingly positive connotations, compared with more neutral terms such as *group* (Delanty 2010; Gold 2005).

Like the sociological and anthropological notions of community, speech communities have traditionally been thought of as linked to a geographical space. Place-based speech communities may range in size from family, village, or city, to nation or larger. Anderson (1983) introduced the term *imagined community* to describe nation-states as socially constructed, since in any community larger than a few thousand people, it is impossible to know all the members personally. This social construction of community not only imagines who is included, but who is excluded. However, Amit and Rapport (2002) argue that place-based definitions poorly reflect transnational communities in a globalized world. Many communities do not live in traditional close-knit groupings (Gumperz and Cook-Gumperz 1982; Milroy 1987) or in the same place; Milburn (2015: 1431) describes this as "[p]erhaps the main challenge to the utility of the term speech communities". Sociolinguists increasingly study transcultural flows: the movements, changes, and reuses of cultural forms in disparate contexts (Pennycook 2007). Many people migrate to cities or diasporas, but may still view themselves as members of a community. Morgan (2005: 4) argues that "[speech community] cannot be defined by static physical location

since membership can be experienced as part of a nation-state, neighborhood, village, club, compound, on-line chat room, religious institution, and so on". Morgan (2014: 109) extends the definition of speech community to discourse styles and communicative norms found in forums, such as the African-American online community 'Black Twitter'.

The place-based definition of community (and speech community) is thus increasingly being expanded to transnational, virtual, and social media communities. These combine the notion of *discourse community* (Swales 1990) with community of practice (Section 7): communication is primarily written, and members "share common values, beliefs, or behaviors" − the third condition in Neal (2012: n.p.). Such phenomena were foreseen by Gumperz (1962: 31) in his definition of linguistic community, which "may consist of small groups bound together by face-to-face contact or may cover large regions, depending on the level of abstraction we wish to achieve".

There is a tendency in linguistics to talk about 'the community' as an idealized homogenous unit, as in "the language attitude of the community itself" (UNESCO 2003: 13). Whaley (2011: 341) illustrates how the very notion of community may be seen as an imposed construct. In the case of the Oroqen in Inner Mongolia, he observes that

> [i]t is not a simple matter to identify a bounded group of people that could be called "the community". Most obviously, clusters of ethnically Oroqen people are distributed over 10 different locations […]. Even within a single location, identifying the Oroqen community is not a straightforward matter. There has been extensive inter-marriage with other ethnic groups. Is the Han man who marries an Oroqen woman part of the Oroqen community? Are their children?

6. Speech

Beyond debates about shared codes versus shared norms, the term *speech* also reflects ideologies about communication. The *Oxford English Dictionary* defines 'speech' primarily as "[t]he expression of or the ability to express thoughts and feelings by articulate sounds", giving the example "he was born deaf and without the power of speech" (www.oxforddictionaries.com/definition/english/speech, last access on 28 August 2018). Likewise, Merriam-Webster's primary definition of 'speech' is "the communication or expression of thoughts in spoken words" (http: //www.merriam-webster.com/dictionary/speech, last access on 28 August 2018). These definitions (especially the OED example) seem to expressly exclude sign languages, despite the existence of speech communities where signing is one communication option open to all members (Zeshan and de Vos 2012). The development of the term speech community predated recognition of sign languages as full-fledged modes of communication. For this reason alone, Gumperz' term 'linguistic community', or the more recent 'community of practice', might be deemed a more inclusive term.

7. Beyond speech communities: communities of practice

As Milburn (2004: 411) points out, "[f]rom Hymes (1962) onward, communication scholars, anthropologists, linguists, sociolinguists, and scholars in ethnic studies have

not only used the term speech communities, but have extended its significance". But has it outlived its usefulness, as Milburn suggests?

Bucholtz (1999: 207) formulated six "ways in which the speech community has been an inadequate model for work on language and gender", some of which have already been discussed:

(a) tendency to take language as central,
(b) emphasis on consensus as the organizing principle of community,
(c) preference for studying central members of the community over those at the margins,
(d) focus on the group at the expense of individuals,
(e) view of identity as a set of static categories, and
(f) valorization of researchers' interpretations over participants' own understandings of their practices.

Some of these criticisms are inevitable, e.g. (a) the centrality of language, given the focus of the fields of sociolinguistics (a subfield of linguistics, the study of language) and linguistic anthropology on language in society (d). The focus on consensual and hegemonic norms (b and c) has already been problematized in this chapter. The issue at stake is whether to discard the concept of speech community as outmoded, or whether to try to incorporate new developments and research paradigms. One of these is the postmodern turn in interpretations of language and identity, which are no longer viewed as "a set of static categories" (point [e] in Bucholtz [1999: 207]), which is sometimes associated with variationist sociolinguistics.

The alternative term *communities of practice*, coined by Lave and Wenger (1991) but developed with regard to language by Eckert (2006, 2012), is increasingly popular. It is based on social practice theory and Bourdieu's (1991) notion of *habitus*, and according to Ortner (2006: 2) aims "to conceptualize the articulations between the practices of social actors 'on the ground' and the big 'structures' and 'systems' that both constrain those practices and yet are ultimately susceptible to being transformed by them". Eckert proposes a theory of language variation as social practice that sees speakers constituting social categories and actively constructing the social meaning of variation. This approach is most clearly articulated in her ethnography of Belten High, a school in Detroit, USA (Eckert 2000).

According to Eckert (2006: 109), a community of practice

> identifies a social grouping not in virtue of shared abstract characteristics (e.g. class, gender) or simple co-presence (e.g. neighborhood, workplace), but in virtue of shared practice. In the course of regular joint activity, a community of practice develops ways of doing things, views, values, power relations, ways of talking.

It could be argued that this is not so different from the shared norms associated with speech communities. The key feature of community of practice which distinguishes it from speech community seems to be "mutual engagement in some common endeavor" (Eckert n.d.), or shared goals rather than shared norms. The community of practice model encourages a more emic interpretation of social categories and linguistic styles, as used in the construction of social meaning in specific contexts. This move to more participant-focused interpretations addresses the hubris implied in Bucholtz' criticism (f), but caution is needed to avoid hypervalorizing or essentializing communities. It has also been

argued in this chapter that the pre-ordaining of social categories criticized by Eckert in variationist sociolinguistics (and by extension, in speech community) is not necessarily inherent in the concept of speech community.

The pendulum in sociolinguistics seems to be swinging away from consensus and social norms, and towards agency and individual choices (e.g. Bucholtz 1999; Coulmas 2013), which is associated with a renewed emphasis on ethnographic research. Eckert (n.d.) asserts that "[T]he individual constructs an identity − a sense of place in the social world − through participation in a variety of communities of practice, and in forms of participation in each of those communities". In this respect, a community of practice seems to have similarities to a 'social network' in Milroy's (e.g. 1987) sense, where group association is based on the strength of individuals' social ties.

Le Page and Tabouret-Keller (1985: 182) proposed a hypothesis whereby "[t]he individual creates his [sic] systems of verbal behaviour so as to resemble those common to the group or groups with which he [sic] wishes from time to time to be identified", with four provisos concerning identification, motivation (reinforced by group feedback), access, and ability to modify behavior (this might nowadays be termed 'agency'). Individuals' linguistic choices are here linked to group norms, "with each individual's knowledge of the systems of his groups the lynch-pin on which the shared concept of communal languages or varieties turns" (Le Page and Tabouret-Keller 1985: 182).

Research into relatively stable multilingual speech communities focuses on individual practices: for example, Mu's (2018) study of motivations for linguistic choices in a multilingual village in southwest China. Mendoza-Denton (2011) describes how research into conscious style-shifting has challenged theories based on covert norms: individuals perform and construct identities and communities through speech styles. She postulates that "individuals may belong to more than one sub-community, and in many cases the sub-communities themselves will be defined in terms of networks and practices and will merge, overlap or otherwise coalesce to form even larger communities" (Mendoza-Denton 2011: 186).

Morgan (2005: 4−5) argues that adults often experience multiple communities: one's initial socialization into a speech community may occur within a culture with communicative values that differ from those of other cultures and communities one encounters later in life, which aligns with current representations of identity as fluid and contingent. Milburn (2015), too, argues that an individual may belong to several, overlapping speech communities. It can also be argued that the literature on the sociology of language in the 1970s and 1980s, notwithstanding its group perspective, paid considerable attention to aspects of fluidity, the mutability of group membership, and different degrees of salience of language as a marker of ethno-linguistic identity (e.g. Fishman 1972, 1977; Le Page and Tabouret-Keller 1985).

8. Conclusions

Morgan (2005, 2014) attempts to bring the notion of speech community up to date, incorporating a broader view than older definitions allowed for: "[s]peech communities are groups that share values and attitudes about language use, varieties and practices" (2014: 1). As Points (2015: 363) observes in her review, Morgan's definition "reflects

previous definitions while evolving the notion of speech communities to be in line with current research ideologies". This chapter, too, argues that although the concept of speech community has been criticized for a tendency to essentialize and to privilege monolingualism, these are not necessarily inherent features, and the concept can be considered more flexibly and inclusively. To sum up, as Eckert (2006: 111) recognizes:

> [i]t should be clear that the speech community and the community of practice approaches are both necessary and complementary, and that the value of each depends on having the right abstract categories and finding the communities of practice in which those categories are most salient. In other words, the best analytic process would involve feedback between the two approaches.

In the 21st century, it might be argued that we are no closer to an agreed definition of 'speech community' than we were 50 years ago, yet it continues to be debated. Indeed, the very 'fuzziness' associated with the term 'speech community' may explain its continued use, despite numerous challenges, criticisms, and proposed alternatives. The term is sufficiently broad and useful to be reinvented and reclaimed by each generation to suit new paradigms of research.

9. References

Amit, Vered & Nigel Rapport
2002 *The trouble with community: Anthropological reflections on movement, identity and collectivity*. London: Pluto Press.

Ammon, Ulrich
1995 *Die deutsche Sprache in Deutschland, Österreich und der Schweiz: das Problem der nationalen Varietäten*. Berlin: Walter de Gruyter.

Anderson, Benedict
1983 *Imagined communities: Reflections on the origin and spread of nationalism*. London & New York: Verso.

Austin, Peter K. & Julia Sallabank
2014 Introduction. In Peter K. Austin & Julia Sallabank (eds.), *Endangered languages: Beliefs and ideologies in language documentation and revitalization*, 1–25. Oxford: Oxford University Press/British Academy.

Bloomfield, Leonard
1933 *Language*. New York: Henry Holt.

Bourdieu, Pierre
1991 *Language and symbolic power*. Cambridge: Polity Press.

Bucholtz, Mary
1999 "Why be normal?" Language and identity practices in a community of nerd girls. *Language in Society* 28(2). 203–223.

Childs, Tucker, Jeff Good & Alice Mitchell
2014 Beyond the ancestral code: Towards a model for sociolinguistic language documentation. *Language Documentation and Conservation* 8. 168–191.

Corder, S. Pit
1973 *Introducing applied linguistics*. Harmondsworth: Penguin.

Coulmas, Florian
2013 *Sociolinguistics: The study of speakers' choices*, 2nd edn. Cambridge: Cambridge University Press.

Delanty, Gerard
2010 *Community*, 2nd edn. Abingdon: Routledge.
Di Carlo, Pierpaolo & Jeff Good
2014 What are we trying to preserve? Diversity, change and ideology at the edge of the Cameroonian Grassfields. In Peter K. Austin & Julia Sallabank (eds.), *Endangered languages: Beliefs and ideologies in language documentation and revitalization*, 229–263. Oxford: Oxford University Press/British Academy.
Eckert, Penelope
2000 *Linguistic variation as social practice*. Oxford: Blackwell.
Eckert, Penelope
2006 Communities of practice. In Jacob L. Mey (ed.), *Concise encyclopedia of pragmatics*, 109–113. Amsterdam: Elsevier.
Eckert, Penelope
2012 Three waves of variation study: The emergence of meaning in the study of variation. *Annual Review of Anthropology* 41(1). 87–100.
Eckert, Penelope
No date Communities of practice, style and personae. http://www.stanford.edu/~eckert/csofp.html/ (accessed 28 August 2016).
Endangered Language Alliance
2012 http://elalliance.org/ (accessed 09 January 2018).
Ferguson, Charles F.
1959 Diglossia. *Word* 15(2). 325–340.
Fishman, Joshua A.
1967 Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues* 23(1). 29–38.
Fishman, Joshua A.
1972 *The sociology of language*. Rowley, MA: Newbury House.
Fishman, Joshua A.
1977 Language and ethnicity. In Howard Giles (ed.), *Language, ethnicity and intergroup relations*, 15–58. London/New York: Academic Press.
Fishman, Joshua A.
1991 *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Bristol: Multilingual Matters.
García, Ofelia & Harald Schiffman (with Zeena Zakharia)
2006 Fishmanian sociolinguistics (1949 to the present). In Ofelia García, Rakhmiel Peltz, Harold Schiffman, & Gella Schweid Fishman (eds.), *Language loyalty, continuity and change: Joshua A. Fishman's contributions to international sociolinguistics*, 3–68. Bristol: Multilingual Matters.
Gold, Ann Grodzins
2005 *Conceptualizing community: Anthropological reflections. A background paper for the Collaborative Initiative for Research Ethics in Environmental Health*. Syracuse, NY: South Asia Center, Moynihan Institute of Global Affairs, Syracuse University.
Gumperz, John J.
1962 Types of linguistic communities. *Anthropological Linguistics* 4(1). 28–40. [Reprinted in Fishman, Joshua (ed.). 1968. *Readings in the sociology of language*, 460–472. The Hague: Mouton]
Gumperz, John J.
1971 [1968] The speech community. In John J. Gumperz (ed.), *Language in social groups*, 114–128. Stanford, CT: Stanford University Press.
Gumperz, John J. & Jenny Cook-Gumperz
1982 Introduction: Language and the communication of social identity. In John J. Gumperz (ed.), *Language and social identity*, 1–22. Cambridge: Cambridge University Press.

Hymes, Dell
1964 *Language in culture and society: A reader in linguistics and anthropology.* New York: Harper and Row.
Hymes, Dell
1967 Models of the interaction of language and social life. *Journal of Social Issues* 23(2). 8–28.
Hymes, Dell
1968 The ethnography of speaking. In Joshua A. Fishman (ed.), *Readings in the sociology of language*, 99–138. The Hague: Mouton.
Hymes, Dell
1974 *Foundations of sociolinguistics*. Philadelphia: University of Pennsylvania Press.
Johnstone, Barbara & William Marcellino
2010 Dell Hymes and the ethnography of communication. In Paul Kerswill, Barbara Johnstone & Ruth Wodak (eds.), *The Sage Handbook of Sociolinguistics*, 57–66. London: Sage.
Kerswill, Paul
2001 A dialect with "inner strength"? The perception of nativeness in the Bergen speech community. *Reading Working Papers in Linguistics* 5. 23–49.
Labov, William
1972 *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
Labov, William
2006 [1966] *The social stratification of English in New York City*, 2nd edn. New York: Cambridge University Press [Washington, DC: Centre for Applied Linguistics].
Lave, Jean & Etienne Wenger
1991 *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
Le Page, Robert B. & Andrée Tabouret-Keller
1985 *Acts of identity: Creole-based approaches to language and ethnicity.* Cambridge: Cambridge University Press.
Loveday, Leo J.
1996 *Language contact in Japan: A socio-linguistic history.* Oxford: Oxford University Press.
Lüpke, Friederike
2016 Pure fiction: The interplay of indexical and essentialist language ideologies and heterogeneous practices. A view from Agnack. In Mandana Seyfeddinipur (ed.), *African language documentation: New data, methods and approaches*, 8–39. Hawai'i: University of Hawai'i Press.
Lyons, John (ed.)
1970 *New directions in linguistics*. London: Penguin.
Madera, Mónica
1996 Speech Community. In Hans Goebl, Peter H. Nelde, Zdenek Stary & Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact Linguistics / Linguistique de contact*, vol. 1, 169–175. Berlin: De Gruyter.
Mendoza-Denton, Norma
2011 Individuals and communities. In Paul Kerswill, Barbara Johnstone & Ruth Wodak (eds.), *The Sage handbook of sociolinguistics*, 181–191. London: Sage.
Migge, Bettina & Isabelle Léglise
2013 *Exploring language in a multilingual context: Variation, interaction and ideology in language documentation.* Cambridge: Cambridge University Press.
Milburn, Trudy
2004 Speech community: Reflections upon communication. In Pamela J. Kalbfleisch (ed.), *Communication Yearbook*, vol. 28, 411–441. Washington, DC: International Communication Association.

Milburn, Trudy
2015 Speech Community. In Karen Tracy, Cornelia Ilie & Todd Sandel (eds.), *International encyclopedia of language and social interaction*, vol. 3, 1427–1432. Washington, DC: John Wiley & Sons with the International Communication Association.
Milroy, Lesley
1987 *Language and social networks.* Oxford: Blackwell.
Milroy, Lesley & James Milroy
1992 Social network and social class: Towards an integrated sociolinguistic model. *Language in Society* 21(1). 1–26.
Morgan, Marcyliena
2005 Speech community. In Alessandro Duranti (ed.), *Companion to linguistic anthropology*, 3–22. Oxford: Wiley-Blackwell.
Morgan, Marcyliena
2014 *Speech communities*. Cambridge: Cambridge University Press.
Mu, Yingying
2018 *Languages, ideologies and identities in Mengguang, a multilingual village in south-western China*. London, UK: SOAS University of London PhD thesis.
Neal, Zachary
2012 Community. In Jeffrey Manza (ed.), *Sociology: Oxford bibliographies*. New York: Oxford University Press. http://www.oxfordbibliographies.com/browse?module_0=obo-9780199756384/ (accessed 29 August 2016). DOI: 10.1093/obo/9780199756384-0080.
Nieldzielski, Nancy A. & Dennis R. Preston
2003 *Folk linguistics.* Berlin: Mouton de Gruyter.
Ortner, Sherry B.
2006 *Anthropology and social theory: Culture, power, and the acting subject*. Duke University Press.
Patrick, Peter L.
2003 The speech community. In J. K. Chambers, Peter Trudgill & Natalie Schilling (eds.), *Handbook of language variation and change*, 573–598. Oxford: Wiley-Blackwell.
Pennycook, Alastair
2007 *Global Englishes and transcultural flows.* London & New York: Routledge.
Pennycook, Alastair
2010 *Language as a local practice.* London: Routledge.
Points, Kathleen
2015 Review of Morgan, Marcyliena H. 2014: *Speech communities. Journal of Linguistic Anthropology* 25(3). 362–364.
Rickford, John R.
1993 Concord and conflict in the speech community. Revised version of Rickford (1986) Concord and contrast in the characterization of the speech community. *Sheffield Working Papers in Language and Linguistics* 3. 87–119.
Schieffelin, Bambi B., Kathryn A. Woolard & Paul V. Kroskrity
1998 *Language ideologies: Practice and theory.* Oxford: Oxford University Press.
Schiffman, Harold F.
1996 *Linguistic culture and language policy.* London: Routledge.
Silverstein, Michael
1996 Monoglot "standard" in America: Standardization and metaphors of linguistic hegemony. In Donald Brenneis & Ronald K. S. Macaulay (eds.), *The matrix of language: Contemporary linguistic anthropology*, 284–306. Boulder: Westview Press.
Spolsky, Bernard
2004 *Language policy.* Cambridge: Cambridge University Press.
Steger, Manfred B.
2003 *Globalization.* Oxford: Oxford University Press.

Swales, John M.
1990 The concept of discourse communities. In John M. Swales (ed.), *Genre analysis: English in academic and research settings*, 21–32. Cambridge: Cambridge University Press.
UNESCO Ad Hoc Expert Group on Endangered Languages
2003 *Language vitality and endangerment: By way of introduction.* http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/pdf/Language_vitality_and_endangerment_EN.pdf/ (accessed 20 March 2018).
Wardhaugh, Ralph
2006 *The study of sociolinguistics*, 5th edn. Oxford: Wiley-Blackwell.
Weinreich, Uriel
1953 *Languages in contact: Findings and problems*. New York: Publications of the Linguistic Circle of New York.
Weisgerber, Leo
1967 *Die Sprachgemeinschaft als Gegenstand sprachwissenschaftlicher Forschung*. Berlin: Springer.
Whaley, Lindsay J.
2011 Some ways to endanger an endangered language project. *Language and Education* 25(4). 339–348.
Winford, Donald
2003 *An introduction to contact linguistics*. Oxford: Blackwell.
Zeshan, Ulrike & Connie de Vos
2012 *Sign languages in village communities: Anthropological and linguistic insights*. Berlin: Walter de Gruyter.

Julia Sallabank, London (UK)

27. Urban and rural language contact

1. Introduction
2. Contesting the urban-rural dichotomy
3. Background to research on urban and rural language contact
4. Language contact phenomena in Africa
5. Consequences of rural-urban language contact
6. Rural-urban language contact and identity
7. Conclusion
8. References

1. Introduction

Language contact, rural or urban, has been the norm in communities around the world, and is even more so in today's globalized world (Coupland 2013), where people are constantly on the move in search of trade, educational, or employment opportunities, or as a result of voluntary or involuntary migrations (Canagarajah 2017). But as Wei (2013: 30) remarks, one does not have to move to a different place to come into contact with people speaking a different language, as there are plenty of opportunities for language contact in the same community, the same neighborhood, or even the same family.

https://doi.org/10.1515/9783110435351-027

In this chapter, we survey language contact phenomena in urban and rural contexts, with a focus on Africa, for we are more familiar with cases of language contact in this region than with comparable cases elsewhere (see e.g. Smakman and Heinrich [2018] for a discussion on language contact in urban areas around the world, Bortoni-Ricardo's [2008] work on rural-urban language contact in Brazil, and Kerswill's [1994] seminal work on rural speech in urban Norway). The survey provides the reader with a better understanding of how language contact phenomena emerge and develop, and how they are used and evaluated in Africa's rural and urban settings. Furthermore, it explores the similarities and differences, if any, between urban and rural language contact, the relevance of such a distinction in the 21st century, and the challenge that ongoing phenomena of migration and globalization pose to the rural-urban dichotomy. We start with a review of the debated dichotomies between both rural and urban spaces, and rural and urban language contact, which have characterized earlier literature on language contact. Next, we describe earlier research into rural and urban language contact to provide the historical background against which language contact phenomena in Africa, the focus of Section 4, will be discussed. Section 5 briefly discusses one of the consequences of rural-urban language contact: language shift for speakers of minority languages. In Section 6, we consider language practices and identity construction in rural-urban language contact in Africa. We conclude with a summary of the key issues covered in the chapter. Our discussion of rural and urban language contact is undergirded not only by research in sociolinguistics, but also by research in cognate fields, especially sociology, geography, and anthropology, where the rural-urban dichotomy has been challenged.

2. Contesting the rural-urban dichotomy

It is impossible to discuss rural-urban language contact without first addressing the complex notions of and the relationship between *rural* and *urban*. Recent studies, particularly in sociology and in social and human geography, argue that there are no universal 'one-size-fits-all' definitions of rural and urban. As such, they challenge the view that the relationship between rural and urban is essentially dichotomous (Champion and Hugo 2004; Halfacree 2004). As Halfacree (2004) argues, the rural-urban relationship should be viewed as gradient and multi-dimensional. Dupont (2004: 177) also problematizes the dualist view of rural-urban in her study of the processes of *peri-urbanization* and *rurbanization* in New Delhi, India. She illustrates rural-urban integration and complexity by pointing to what she calls 'urban villages' − areas where low-income migrants, primarily from rural areas, live because they cannot afford to live in the capital city where they work. Along these lines, Patrick and Budach (2014) argue that because of the increasing flow of people moving back and forth between spaces traditionally conceived as rural or urban, the long-standing rural-urban dichotomy has little descriptive or explanatory utility. Deumert (2013: 55) makes a similar comment in her study of rural-urban migration in South Africa, noting that moving from a rural to an urban area does not necessarily mean cutting off all ties with a prior location; instead, it can involve "a series of interlocking, shuffling movements between the two places". Thus, South African rural-urban migrants are neither entirely rural nor exclusively urban, for they sometimes reside in one, but work in the other; the linguistic resources associated with the

rural home continue to play a role in the urban environment (Deumert 2013). Potts (2010) similarly observes that many rural-urban migrants in Zimbabwe, as well as other African countries, move back and forth between rural and urban areas and, as a result, claim and maintain their memberships and identities in both rural and urban communities. Accordingly, Lang (1986: 120) suggests − and this is the view that will guide our discussion of language contact phenomena in this chapter − that rural and urban be viewed as "denot[ing] opposite ends of the *conceptual continuum* [emphasis ours]". On this conceptual continuum, or what Overbeek and Terluin (2006: 22) describe as the *rurban*, people and communities fall somewhere between the two hypothetical extremes.

3. Background to research on urban and rural language contact

Studies on urban and rural language contact were traditionally conducted in light of the purported dichotomy between urban and rural spaces, which was championed in early social sciences research (Durkheim 1964; Wirth 1938). Bell (1992) remarks that for that research, the 'common wisdom' was that city life and country life were markedly different. A sharp demarcation between urban and rural in social science research influenced early linguistic studies on language contact, especially 19th century dialectology, with its focus on regional varieties (Ellis 1889; Wright 1892). In particular, traditional dialectology − heavily influenced by historical linguistics − was primarily concerned with identifying the geographic boundaries of particular linguistic features (known as *isoglosses*), creating dialect maps and atlases, and providing formal descriptions of unadulterated, or 'pure', dialects. It was believed that such dialects were more prone to change in the cities, where the population is typically more modern, mobile, and educated, and thus affected by the processes of language standardization and contact. Furthermore, women were perceived as being less inclined to use vernacular, and thus were typically excluded as study participants. Therefore, NORMs (non-mobile, old, rural males) were considered to be the carriers and preservers of 'pure' dialects, and their language practices served as the source of data for dialect studies (Chambers and Trudgill 1998). Ironically, traditional dialectology both relied on NORMs, whose somewhat archaic speech patterns did not represent the rest of the community, and simultaneously viewed rural communities as socially and linguistically homogeneous.

In contrast to early dialectology studies, which privileged rural speech, linguistic variation studies have focused on urban speech (Labov 1972; Milroy and Milroy 1978), thus leading to *metro-normativity* (Britain 2012; May 2014), a view that tends to normalize urban varieties while ignoring and/or marginalizing rural ones. Furthermore, the fascination with and focus on urbanity and its concomitant metro-normativity have, perhaps unintentionally, also resulted in reasserting the 'old' rural-urban divide, in which 'rural' is viewed as detached from global processes as well as socioculturally and economically regressive (Woods 2011), and thus as the opposite of 'urban'. Yet, as Britain (2012: 16), drawing on Trudgill (1986), aptly observes, language contact "is insensitive to urban or rural locale. It just happens that such contact is perhaps more vividly and frequently experienced and perceived in cities, consequently helping to explain why it and urbanization may have become confused. But, radical dialect contact can occur in highly rural areas as well as in cities and the outcomes are typologically the same". It follows that

new linguistic varieties and language contact phenomena, including koineization, dialect leveling, and interdialect, occur in both rural and urban language contact areas (Britain 2012: 22).

Bearing the foregoing in mind, it is impossible to draw a clear line between rural language contact and urban language contact, particularly in today's increasingly globalized world. Instead, as already noted, we consider rural-urban language contact to be a conceptual continuum, in the sense of Lang (1986). This continuum is neither entirely urban nor entirely rural, since both urban and rural do − as a result of continual back and forth migration flows, globalization, and advances in technology − mutually influence each other. Thus, urban language contact, just like rural language contact, is a fluid and dynamic phenomenon, as research on language contact in African settings demonstrates (Mensah 2016; Nassenstein 2016; Nassenstein and Hollington 2015). Such research not only offers in-depth explorations of African urban youths' repertoires (their features and functions), but also suggests that youth language varieties and urban vernaculars in general, though typically originating and developing in the cities, eventually transfer to the rural areas, where they undergo further changes and developments (Kioko 2015; Nassenstein 2016).

4. Language contact phenomena in Africa

The extant literature indicates that when scholars speak of language contact, they are really talking about people using different languages when coming into contact with one another (Brenzinger 1997; Wei 2013). Contact between languages, in urban or rural contexts, usually leads to what Haugen (1972) termed *interlingual contagion*, that is, the impact that the languages involved have on one another. This impact may manifest itself not only in such widely researched language contact phenomena as borrowing, codeswitching, diglossia, pidginization, creolization, language shift, language death, and language conflict (Kamwangamalu 2008; Potowski 2013), but also in related phenomena such as colingualism (Schell 2008), passing, refusal and crossing (Rampton 1995, 2017), translanguaging (García 2009; García and Wei 2014), codemeshing (Canagarajah 2011), and mixed, increasingly multilingual, varieties such as Sheng in Kenya (Kioko 2015), Tsotsitaal in South Africa (Brookes 2014), and Pidgin English in Nigeria (Gbogi 2016). In the next three sections, we describe some of these lesser-studied language contact phenomena and mixed varieties in Africa. Our goal is to provide the reader with a better understanding of how these phenomena emerge and develop, and how they are used and evaluated in Africa's rural-urban setting continuum.

4.1. Crossing, refusal, passing

The language contact phenomena discussed in this section are arguably more prevalent in urban settings than in rural settings due to the dense multilingual nature of the former. Rampton (1995: 280) describes *crossing* as "code alternation by people who are not accepted members of the group associated with the second language they employ. It is concerned with switching into languages that are not generally thought to belong to

you". *Code alternation*, says Gumperz (1982), is a contextualization cue, or a linguistic (e.g. word choice) or paralinguistic (e.g. intonation) feature, that speakers and listeners use to signal and interpret meaning in an interaction. Viewed from this perspective, crossing is similar to the widely studied phenomenon of *codeswitching*, for both phenomena are contextualization cues. However, the two language contact phenomena do differ in many respects. Crossing is, first and foremost, an outgroup phenomenon, but as Woolard (1988: 69–70) notes, codeswitching "is an in-group phenomenon, restricted to those who share the same expectations and rules of interpretation for the use of the two languages". In crossing, code alternation bears the distinctive characteristic of being always marked in the sense of Myers-Scotton (1997), that is, it always entails use of the least expected code or language. Also, crossing refers "to practices which involve *a sharp sense of social or ethnic boundary transgression*" (Rampton 2017: 10, italics in the original). These practices, we further note, can be negatively sanctioned if they take place at the wrong time, in the wrong place, and/or in the wrong company. Also noteworthy to the distinction between crossing and codeswitching is that the latter presupposes no violation of the norms that govern language use in the community of which the participants are members. In crossing, however, the speaker may choose to challenge these norms, for "crossing can raise serious issues of legitimacy and entitlement, potentially provoking recipients to respond with the political challenge, 'by what right do you use that language?'" (Rampton 2017: 10).

Related to crossing are such language contact phenomena as *refusal* and *passing*. Rampton (1995) describes refusal as a practice where individuals avoid claiming a language that is not their own, a phenomenon that is associated with passing. The latter, says Rampton (1995: 287), is a practice where, in order to avoid any talk that draws attention to their use of an out-group code, people pretend that the out-group code is actually part of their own inheritance (Rampton 1995: 287). According to this practice, people attempt to become native speakers of the language they cross into, that is, "they lie about their place and language of origin" (Rampton 1995: 287–288). For example, in a study of rural Zimbabwean migrants to South African cities, Maphosa (2009: 29) reports that Zimbabweans from the Matabeleland "pass" and thus avoid arrests and deportations because "they easily master South Africa languages such as Zulu, which [is similar] to Ndebele, spoken in some parts of Matabeleland". Laitin (1993: 59) remarks that the choices that such immigrants make to integrate into the host community can be explained from the perspective of what he calls "a competitive assimilation game": despite the loyalty they might have for their home language, working-class or unemployed immigrants have a strategic incentive to assimilate the language of their new home in order to be able to compete for middle-class jobs.

4.2. Colingualism

In urban or rural Africa, as elsewhere in postcolonial settings, a number of contextual factors (e.g. topic, setting, and/or participants) determine daily linguistic practices of people who share more than one language. Schell (2008) raises the question of what to call people who, even if they share two languages, L1 and L2, where L1 is the first learned language (often referred to as mother tongue) and L2 is the second language, tend to communicate via the medium of their L2, which in Africa may be a former

colonial language (English, French, Portuguese, or Spanish), or an indigenous lingua franca (Swahili in East Africa, Zulu in South Africa, or Wolof in Senegal). Schell (2008: 120) proposes that such people be called *colinguals*. Colingualism, we argue, is a by-product of bilingualism involving an indigenous language, and either a local lingua franca or a former colonial language, as noted above. To our knowledge, colingualism, let alone its consequences − to which we will return later − for minority language maintenance, has hardly been investigated in the African context. Thus, it remains a fertile ground for future research on language contact phenomena on the continent.

4.3. From urban to rural and back: Africa's mixed linguistic varieties and their trajectories

The continent's randomly selected mixed linguistic varieties to be described here include Tsotsitaal, Sheng, and Nigerian Pidgin English. Our goal, again, is to discuss the similarities and differences, if any, between the trajectories of their development and use. It is important to point out that each of the listed mixed linguistic varieties, as any linguistic variety in general, is better understood as a continuum of evolving varieties and not as a monolithic unit (see Kioko 2015). The literature has used various labels to describe such varieties, among them polylingual languaging (Jørgensen 2008), metrolingualism (Pennycook and Otsuji 2015), contemporary urban vernaculars (Rampton 2011), and codemeshing (Canagarajah 2011). These widespread mixed linguistic varieties are typically linked to urban living and the rural-urban language contact continuum. They are, as Mensah (2016: 1) describes them, a byproduct of the combined forces of urbanization, migration, multilingualism, globalization, and transcultural contacts. Often referred to as urban youth language practices or styles, the selected mixed linguistic varieties display similarities both in terms of their linguistic structures and their social functions. Structurally, the literature lists the following as common strategies, among others, employed in their development and emergence: use of taboo forms and expressions; vulgar terms of address; slang and metaphors; phonological, morphological, and semantic manipulations; deviant vocabulary; codeswitching; idiomatic practices; metathesis; and appropriation of loan words (Mensah 2016; Mugaddam 2009; Nassenstein 2016).

The functions of Africa's mixed linguistic varieties are determined by the sociolinguistic context in which they have emerged. For example, in his study of Yabacrane − a Kiswahili-based urban youth language in Goma in the eastern Democratic Republic of the Congo (DRC) − Nassenstein (2016) notes that this variety reveals salient features that do not necessarily coincide with characteristics typical of other youth language practices in Africa. He explains that "the gangster image and violent behavior, aggravated by the prevailing conflict zone in Eastern DR Congo, as well as a tough 'street image', are the central points of reference for Yabacrane speakers" (Nassenstein 2016: 235). These features, it is noted further, do not apply to all African mixed linguistic varieties, nor are they limited to a certain stratum of the population. The varieties are also spoken by university students, who are not typically associated with criminal activity or a 'tough street' image. African mixed linguistic varieties do, however, serve some common functions for their speakers, such as indexing identity, enhancing solidarity, fostering group integration, reinforcing anti-establishment and resistance behavior, maintaining

secrecy, projecting competitive in-group self-expression, protesting social inequality, and so on (Mensah 2016; Nassenstein and Hollington 2015).

Although Sheng, Tsotsitaal, and Nigerian Pidgin English are connected to rural-urban migration in bilingual and multilingual contexts, each of them followed a unique path. For instance, Tsotsitaal (an Afrikaans-based, and increasingly Bantu-based, set of mixed varieties) traces its origins to rural-urban migrant labor and forced migration (Morris 2010). With time (1940−1950), it came to index covert, and often hypermasculine, prestige linked to "fighting skills, independence, street wisdom, feats of daring, law-breaking, clothing style […] and success with women" (Glaser 2000: 4); thus, it became associated with toughness of the streets and street-smarts, a type of gang 'argot'. As such, it has become closely linked to urban, and often dangerous, youth culture. It evolved even further, and now signifies "cosmopolitanism internal to township life" (Morris 2010: 105) and "the high life of the city − the urban, the cool, the hip, and the sophisticated" (Childs 1996: 342).

A trajectory similar to that of Tsotsitaal can be ascribed to Sheng, a Swahili-based, and now increasingly English-based, mixed variety in Kenya. From a linguistic variety associated with the poorest areas of Nairobi that numerous migrants − typically from rural areas and with diverse ethno-linguistic backgrounds − call home (Kang'ethe-Iraki 2004), Sheng has come to be associated with young people in general, a type of the 'urban cool'. Bosire (2006: 192) describes this variety as "a way to break away from the old fraternities that put particular ethnic communities in particular neighborhoods/ 'estates' and give them a global urban ethnicity, the urbanite: sophisticated, street smart, new generation, tough". In a study that we conducted in 2013 in Kenya, which is part of a larger research project (the DoD/Howard University cooperative agreement 59530-MA-PIR) and from which we also draw in the remainder of the chapter, many of the participants indicated that if young people did not speak Sheng, they would be viewed as 'local', not cosmopolitan, and definitely as 'rural' (Kamwangamalu and Tovares 2016; Tovares and Kamwangamalu 2017). One participant, whom we call Alvin (all participants' names have been anonymized), describes this as follows: "Were you born in the rural [area]? Did you live there [so] that you can't speak Sheng?" Another participant, Amanda, stresses that Sheng indexes one's identity as a young person: "If you hear a person, even if you don't see them, if you hear it is Sheng being spoken, you know that's a youth". Not only does Sheng − as a type of a stylistic choice − index young, urban, modern, and hip identities, but it may also index one's urban beginnings. In addition, Sheng is increasingly viewed as a bridge between the poor and the rich, and as a common denominator among young people of various ethnicities. Music, the media, and popular culture reflect and co-construct the acceptance of Sheng in Kenyan society.

Like the mixed language varieties discussed above, some pidgin languages are also inextricably linked to the rural-urban language contact continuum. One such language is Nigerian Pidgin English. Onumajuru and Onumajuru (2016) observe that rural migrants have contributed to the heterogeneous linguistic situation in Nigerian cities. They suggest that, since the languages spoken in the cities are mutually unintelligible, urban dwellers use either English or − even more so − Nigerian Pidgin English. As Dueber's (2005) study of Nigerian Pidgin English shows, in Lagos − and other large Nigerian cities − the linguistic situation can be described as triglossic: English, Nigerian Pidgin English, and indigenous languages. According to Gbogi (2016: 172), with the legacy of English as the former colonial language and indigenous languages' ties to different ethnicities

and tribes, Nigerian Pidgin English is perceived as a 'neutral' language devoid of political or ethnic overtones. Although initially a lingua franca used by linguistically and ethnically diverse Nigerians, Nigerian Pidgin English has become a language of everyday communication in urban settings, from markets to university campuses, and "is used to perform identities (ranging from cultural, political, bi/multilingual, to global) and sustain the in-groupness of a growing urban youth subculture in Nigeria" (Gbogi 2016: 188). Furthermore, Gbogi (2016) argues that Nigerian Pidgin English, similar to some mixed varieties, such as Sheng and Tsotsitaal (as discussed above), is not only enjoying popularity in the urban areas, and progressively in rural areas, but is also becoming the first or second language of an increasing number of people (more than 30 million).

5. Consequences of rural-urban language contact

The literature indicates that the language contact phenomena described in the previous sections have consequences for speakers of minority languages in both rural as well as urban spaces. One such consequence is language shift from a minority language to a local lingua franca, or to a former colonial language. *Language shift*, says Fasold (1984), is a gradual process in which a speech community, for one reason or another, gives up its language and adopts a new one. The starting point of language shift is bilingualism − often accompanied by diglossia − as a stage on the way to monolingualism in a new language (Romaine 1994). For instance, Kioko (2015: 130) notes the following regarding the impact of Sheng on local vernaculars in Kenya:

> Given the rate of rural-urban migration in Kenya today, a sizable percentage of children are born and socialized in multi-ethnic urban settings where daily interactions and communication are devoid of mother tongue. Therefore, most urban-socialized children and youths lack knowledge of their mother tongue(s) [ethnic language varieties] and are thus increasingly identifying with Sheng as a first language.

In a related study, Kamwangamalu (2013) reports that African migrants in the Washington metropolitan area are mostly colinguals, that is, they tend to communicate through the medium of a former colonial language (in this case, French), which they learned prior to migrating to the United States. Consequently, not only migrants, but also their children, are gradually shifting from their heritage language(s) to a former colonial language. For the children, another round of language shift appears to be on the way, the shift from the transplanted language (French) to English. This is due, in part, to what Tuominen (1999) has termed *utility-maximization*, that is, the costs and benefits of raising children in a community's ethnic language, especially if the language has no status in the surrounding society. It seems that, as Clyne and Kipp (2006) observe, the perceived cost-benefit analysis tips the balance in favor of a decline in use of the heritage language. In this regard, Paulston (1994: 9) remarks that "ethnic groups within a modern nation-state, given opportunity and incentive, typically shift to the language of the [economically] dominant group". The foregoing discussion also points to a link between language and identity, the focus of the next (and final) section.

6. Rural-urban language contact and identity

Individuals in language contact encounters, rural or urban, use a wide range of linguistic resources to construct their social identities, to question, negotiate, and/or (re)evaluate their and others' identities, language attitudes, and linguistic practices, and to determine whether to converge with, or diverge from, such practices (Gumperz 1982; Kamwangamalu 1992; Zhu 2017). More specifically, we want to determine which aspects of identity, as highlighted earlier with respect to Sheng and Tsotsitaal (e.g. prestige, sophistication, gender, cosmopolitanism, localness, in-groupness, urbanite, youth, etc.), become more salient in urban and rural language contact. To do so, we consider a comparable case study by Hachimi (2012) of two urban dialects in Morocco: the *urban* dialect of Casablanca (a set of linguistic varieties greatly influenced by rural-urban migration, colonization, and industrialization), and the *urbane* linguistic variety of the city of Fez, termed Fessi (a traditionally prestigious dialect linked to the elite). Hachimi (2012) explores the dynamic nature of urban, and by extension rural, linguistic contacts in relation to identity. Specifically, she focuses on how individuals negotiate and project identities in contact situations against the backdrop of the histories and trajectories of the afore-named urban dialects. Whilst language contact studies tend to center on "socially disenfranchised rural groups who bring along 'stigmatized' dialects to the host city" (Hachimi 2012: 322), her work zeros in on the speakers of the prestigious Fessi dialect who migrate from Fez to Casablanca, the center of new urban varieties. Hachimi's (2012) analysis indicates that the older generation continues to link the Fessi dialect to prestige, politeness, and sophistication, but the younger, especially Casablanca-born, Fessis favor the Casablancan dialect. At the same time, this shift − Hachimi observes − is not toward valorizing the rural-linked features of the Casablancan dialect, but rather toward valorizing those that index the covert prestige of 'street smarts'. In this way, the ever-evolving Casablancan varieties are similar to other urban varieties, including Tsotsitaal and Sheng.

Although Tsotsitaal and Sheng are traditionally associated with urban young men, these varieties, and Sheng in particular, are increasingly being used by young women. We document elsewhere (Tovares and Kamwangamalu 2017) that both male and female students in Eldoret, Kenya, report using Sheng with their friends and siblings. For instance, one female participant, whom we call Donna, states that she and her siblings use Sheng to separate themselves from their parents, who "will not understand Sheng". In addition to being used by young men and women, Sheng is gradually crossing generational boundaries. For instance, Tim (another participant in the study) indicates that, because Sheng is used in the media and by a number of celebrities, it is "now being understood even by the parents who are not supposed to get Sheng. So it is a language that is becoming known to everyone". While Sheng is showing signs of growing into an urban and rural, and a somewhat age- and gender-neutral, way of speaking, Tsotsitaal remains strongly − though not exclusively − associated with young urban males, and is often mobilized to index hypermasculinity. Brookes (2014) argues that Tsotsitaal is not only linked to 'street smart' youth, but even more so to gender: it is a 'male youth talk'. She goes on to suggest that young men use Tsotsitaal to "negotiate friendships and social alignments, celebrate their masculinity, and demonstrate streetwise township identities by engaging in verbal exchanges that differentiate one another in terms of linguistic skill, status, and authenticity" (Brookes 2014: 361). In sum, urban linguistic varieties are typically linked to cosmopolitanism, trendiness, and youth; however, rural varieties,

and the ethnic languages that influence them, are typically associated with connection to one's family roots.

7. Conclusion

In this chapter, we discussed the outcomes of urban and rural language contact, with a focus on lesser-studied phenomena, especially crossing, passing, refusal, and colingualism, as well as the mixed varieties Sheng, Tsositaal, and Nigerian Pidgin English. These phenomena occur in both Africa's urban and rural settings, where the dichotomy between the two settings has increasingly become porous as a result of constant back-and-forth migration flows and of globalization. As noted earlier, language contact, rural or urban, may result in language shift or loss, especially for speakers of minority languages. This area − language shift or loss, along with language practices and identity construction in social spaces − remains fertile ground for future research on the rural-urban language contact continuum in Africa.

8. References

Bell, Michael M.
1992 Fruit of difference: The rural-urban continuum as a system of identity. *Rural Sociology* 57(1). 65−82.

Bortoni-Ricardo, Stella M.
2008 *The urbanization of rural dialect speakers: A sociolinguistic study in Brazil*. Cambridge: Cambridge University Press.

Bosire, Mokaya
2006 Hybrid languages: The case of Sheng. In Olaoba F. Arasanyin & Michael A. Pemberton (eds.), *Selected proceedings of the 36th annual conference on African linguistics*, 185−193. Somerville, MA: Cascadilla Proceedings Project.

Brenzinger, Matthias
1997 Language contact and language displacement. In Florian Coulmas (ed.), *The handbook of sociolinguistics*, 273−284. Malden, MA: Blackwell Publishers.

Britain, David
2012 Countering the urbanist agenda in variationist sociolinguistics: Dialect contact, demographic change and the rural-urban dichotomy. In Sandra Hansen, Christian Schwartz, Philipp Stoeckle & Tobias Streck (eds.), *Dialectological and folk dialectological concepts of space*, 12−30. Berlin: Mouton de Gruyter.

Brookes, Heather
2014 Urban youth languages in South Africa: A case study of Tsotsitaal in a South African township. *Anthropological Linguistics* 56(3−4). 356−388.

Canagarajah, A. Suresh
2011 Codemeshing in academic writing: Identifying teachable strategies of translanguaging. *Modern Language Journal* 95. 401−417.

Canagarajah, A. Suresh (ed.)
2017 *Routledge handbook of migration and language*. New York: Routledge.

Chambers, Jack K. & Peter Trudgill
1998 *Dialectology*, 2nd edn. Cambridge: Cambridge University Press.

Champion, Tony & Graeme Hugo
2004 Introduction: Moving beyond the urban-rural dichotomy. In Tony Champion & Graeme Hugo (eds.), *New forms of urbanization: Beyond the urban-rural dichotomy*, 3–24. Burlington, VT: Ashgate.
Childs, George T.
1996 The status of Isicamtho, an Nguni-based urban variety of Soweto. In Arthur K. Spears & Donald Winford (eds.), *The structure and status of pidgins and creoles*, 341–370. Amsterdam: John Benjamins.
Clyne, Michael & Sandra Kipp
2006 Australia's community languages. *International Journal of the Sociology of Language* 180. 7–21.
Coupland, Nikolas (ed.)
2013 *The handbook of language and globalization*. Malden, MA: Wiley-Blackwell.
Deumert, Ana
2013 Xhosa in town (revisited): Space, place and language. *International Journal of the Sociology of Language* 222. 51–75.
Dueber, Dagmar
2005 *Pidgin English in Nigeria: Language contact, variation and change in an African urban setting.* London: Battlebridge.
Dupont, Véronique
2004 Urban development and redistribution in Delhi: Implications for categorizing population. In Tony Champion & Graeme Hugo (eds.), *New forms of urbanization: Beyond the urban-rural dichotomy*, 171–190. Burlington, VT: Ashgate.
Durkheim, Emile
1964 *Division of labor in society*. New York: Free Press.
Ellis, Alexander John
1889 *On early English pronunciation*. London: Trübner & Co.
Fasold, Ralph
1984 *The sociolinguistics of society.* Oxford: Blackwell.
García, Ofelia
2009 *Bilingual education in the 21st century: A global perspective*. Malden, MA: Blackwell.
García, Ofelia & Li Wei
2014 *Translanguaging, language, bilingualism and education.* New York: Palgrave MacMillan.
Gbogi, Michael Tosin
2016 Language, identity, and urban youth culture: Nigerian hip hop music as an exemplar. *Pragmatics* 26(2). 171–195.
Glaser, Clive
2000 *Bo-Tsotsi: The youth gangs of Soweto, 1935–1976.* Cape Town: David Philip.
Gumperz, John
1982 *Discourse strategies*. Cambridge: Cambridge University Press.
Hachimi, Atiqa
2012 The urban and the urbane: Identities, language ideologies, and Arabic dialects in Morocco. *Language in Society* 41(3). 321–341.
Halfacree, Keith
2004 Rethinking 'rurality'. In Tony Champion & Graeme Hugo (eds.), *New forms of urbanization: Beyond the urban-rural dichotomy*, 287–304. Burlington, VT: Ashgate.
Haugen, Einar
1972 The stigmata of bilingualism. In Anwar Dil (ed.), *The ecology of language*, 307–344. Stanford: Stanford University Press.

Jørgensen, J. Normann
2008 Poly-lingual languaging around and among children and adolescents. *International Journal of Multiligalism* 5(3). 161–176.
Kamwangamalu, Nkonko M.
1992 Multilingualism and social identity in Singapore. *Journal of Asia-Pacific Communication* 3(1). 33–47.
Kamwangamalu, Nkonko M.
2008 Colinguals among bilinguals: Commentary 6. *World Englishes* 27(2). 285–287.
Kamwangamalu, Nkonko M.
2013 Rural-urban and south-north migrations and language maintenance and shift. *International Journal of the Sociology of Language* 222. 33–49.
Kamwangamalu, Nkonko & Alla Tovares
2016 English in language ideologies, attitudes, and educational practices in Kenya and South Africa. *World Englishes* 35(3). 421–439.
Kang'ethe-Iraki, Frederick
2004 Cognitive efficiency: The Sheng phenomenon in Kenya. *Pragmatics* 14(1). 55–68.
Kerswill, Paul
1994 *Dialects converging: Rural speech in urban Norway*. Oxford: Clarendon Press.
Kioko, Eric M.
2015 Regional varieties and ethnic registers of Sheng. In Nico Nassenstein & Andrea Hollington (eds.), *Youth language practices in Africa and beyond*, 119–148. Berlin & Boston: Walter de Gruyter.
Labov, William
1972 *Language in the inner city*. Philadelphia, PA: University of Pennsylvania Press.
Laitin, David
1993 Migration and language shift in urban India. *International Journal of the Sociology of Language* 103. 57–72.
Lang, Marvel
1986 Redefining urban and rural for the US Census of Population: Assessing the need for alternative approaches. *Urban Geography* 7(2). 118–134.
Maphosa, France
2009 *Rural livelihoods in Zimbabwe: Impact of remittances from South Africa*. Dakar: CODESRIA.
May, Stephen
2014 Contesting metronormativity: Exploring indigenous language dynamism across the urban-rural divide. *Journal of Language, Identity, and Education* 13(4). 229–235.
Mensah, Eyo
2016 The dynamics of youth language in Africa: An introduction. *Sociolinguistic Studies* 10(1–2). 1–14.
Milroy, James & Lesley Milroy
1978 Belfast: Change and variation in an urban vernacular. In Peter Trudgill (ed.), *Sociolinguistic patterns in British English*, 19–36. London: Edward Arnold.
Morris, Rosalind C.
2010 Style, Tsotsi-style, and Tsotsitaal: The histories, aesthetics, and politics of a South African figure. *Social Text* 28(2). 85–112.
Mugaddam, Abdel
2009 Aspects of youth languages in Khartoum. In Matthias Brenzinger & Anne-Maria Fehn (eds.), *Proceedings of the 6th World Congress on African Linguistics*, 87–98. Cologne: Köppe.
Myers-Scotton, Carol
1997 Codeswitching. In Florian Coulmas (ed.), *The handbook of sociolinguistics*, 217–237. Cambridge, MA: Blackwell Publishers.

Nassenstein, Nico
2016 The new urban youth language Yabacrane in Goma (DR Congo). *Sociolinguistic Studies* 10(1–2). 235–259.
Nassenstein, Nico & Andrea Hollington (eds.)
2015 *Youth language practices in Africa and beyond.* Berlin & New York: Mouton de Gruyter.
Onumajuru, Emeka M. & Virginia C. Onumajuru
2016 Formulaic exchange in Igbo. In Ozo-Mekuri Ndimele (ed.), *Studies in Nigerian linguistics*, 121–134. Port Harcourt, Nigeria: LAN Occasional Publications.
Overbeek, Greet & Ida Terluin
2006 Introduction. In Greet Overbeek & Ida Terluin (eds.), *Rural areas under urban pressure: Case studies of rural-urban relationships across Europe*, 21–26. The Hague: LEI Wageningen UR.
Patrick, Donna & Gabriele Budach
2014 Urban-rural dynamics and indigenous urbanization: The case of Inuit language use in Ottawa. *Journal of Language, Identity, and Education* 13(4). 236–253.
Paulston, Christina B.
1994 *Linguistic minorities in multilingual settings: Implications for language policies*. Amsterdam: John Benjamins.
Pennycook, Alastair & Emi Otsuji
2015 *Metrolingualism: Language in the city.* Abingdon: Routledge.
Potowski, Kim
2013 Language maintenance and shift. In Robert Bayley, Richard Cameron & Ceil Lucas (eds.), *The Oxford handbook of sociolinguistics*, 301–320. Oxford: Oxford University Press.
Potts, Deborah
2010 *Circular migration in Zimbabwe and contemporary Sub-Saharan Africa*. Suffolk, UK: James Currey.
Rampton, Ben
1995 *Crossing: Language and ethnicity among adolescents*. London & New York: Longman.
Rampton, Ben
2011 From "multi-ethnic adolescent heteroglossia" to "contemporary urban vernaculars". *Language & Communication* 31(4). 276–294.
Rampton, Ben
2017 *Crossing* thirty years later. *Working Papers in Urban Language & Literacies*. Paper 210. London: Kings College.
Romaine, Suzanne
1994 *Language in society: An introduction to sociolinguistics*. New York: Oxford University Press.
Schell, Martin
2008 Colinguals among bilinguals. *World Englishes* 27(1). 117–130.
Smakman, Dick & Patrick Heinrich (eds.)
2018 *Urban sociolinguistics: The city as a linguistic process and experience*. London and New York: Routledge.
Tovares, Alla & Nkonko M. Kamwangamalu
2017 Migration trajectories: Implications for language proficiencies and identities. In Suresh Canagarajah (ed.), *Routledge handbook of migration and language*, 207–227. New York: Routledge.
Trudgill, Peter
1986 *Dialects in contact*. Oxford: Blackwell.
Tuominen, Anne
1999 Who decides the home language? A look at multilingual families. *International Journal of the Sociology of Language* 140. 59–76.

Wei, Li
2013 Conceptual and methodological issues in bilingualism and multilingualism Research. In Tej Bhatia & William Ritchie (eds.), *The handbook of bilingualism and multilingualism*, 2nd edn., 27–51. Malden, MA: Blackwell Publishing Company.
Wirth, Louis
1938 Urbanism as a way of life. *American Journal of Sociology* 44(1). 1–24.
Woods, Michael
2011 *Rural.* London: Routledge.
Woolard, Kathryn A.
1988 Codeswitching and comedy in Catalonia. In Monica Heller (ed.), *Codeswitching: Anthropological and sociolinguistic perspectives*, 53–76. The Hague: Mouton.
Wright, Joseph
1892 *A grammar of the dialect of Windhill in the West Riding of Yorkshire*. London: Kegan Paul.
Zhu, Hua
2017 New orientations to identity in mobility. In Suresh Canagarajah (ed.), *Routledge handbook of migration and language*, 117–132. New York: Routledge.

Nkonko M. Kamwangamalu, Washington, DC (USA)
Alla V. Tovares, Washington, DC (USA)

28. Globalization and superdiversity

1. Introduction
2. Language contact studies and the denaturalization of language names and language systems
3. On fluid multilingualism and urban vernaculars
4. Conclusion
5. References

1. Introduction

Globalization is all too often treated as a primarily economic phenomenon (Perrons 2004: 35–54; Wade 2001) even though it is in fact a highly multidimensional set of processes that also include the political, technological, and cultural (Giddens 2002: 10; Kennedy 2001: 8). The key characteristic of globalization is, as Giddens (1990: 64) points out, "the intensification of worldwide social relations which link localities in such a way that local happenings are shaped by events occurring many miles away and vice versa". This intensification is facilitated (some might even say 'exacerbated') by developments and advances in communication and transportation technologies which allow individuals and communities to more easily maintain contact with multiple groups across the globe.

As far as language contact is concerned, this means that it is not only people who are mobile and who therefore bring with them their language practices as they move from one place to another. Languages, too, can move, even without speakers, since media technologies also allow for the relatively rapid and widespread dissemination of language

https://doi.org/10.1515/9783110435351-028

resources − in the form of what we might think of as cultural texts − through processes of streaming and downloading as well as, of course, the more traditional media of radio, movies, and television.

Where the movement of people is concerned, mobility no longer necessarily means loss of contact with the home community. The time that a migrant spends living and working in the host community need not result in a reduction in the frequency of interactions with friends and family members back home. There is obviously no single homogeneous entity that we can meaningfully describe as 'the migrant', since the relevant factors that influence a person's decision to migrate are extremely varied. Faist (2000: 37) gives us a sense of just how varied these factors can be, when he notes that they

> […] may be related to improving and securing: wealth (e.g. income), status (e.g. prestigious job), comfort (e.g. better working and living conditions), stimulation (e.g. experience, adventure, and pleasure), autonomy (e.g. high degree of personal freedom), affiliation (e.g. joining friends or family), exit from oppression of all kinds, meaningful life (e.g. improving society), better life for one's children, and morality (e.g. leading a virtuous life for religious reasons). In this view the potential migrant could not only be a worker, a member of a household or a kinship group, but also a voter, a member of ethnic, linguistic, religious, and political groups, a member of a persecuted minority, or a devotee of arts or sports.

The result of all this variation is what has sometimes been called 'superdiversity' (Vertovec 2007). As Blommaert and Rampton (2016: 22) point out,

> [s]uperdiversity is characterized by a tremendous increase in the categories of migrants, not only in terms of nationality, ethnicity, language, and religion, but also in terms of motives, patterns and itineraries of migration, processes of insertion into the labor and housing markets of the host societies, and so on (cf. Vertovec 2010). The predictability of the category of 'migrant' and of his/her sociocultural features has disappeared.

It may have once been feasible to understand the impact of migration on language contact by identifying specific waves of migration patterns. These included particular types of colonization that carry with them distinct implications for social interaction and communicative practices. Possible examples are trade colonization, exploitation colonization, and settlement colonization (Mufwene 2001), and also more recently, postcolonial Englishes (Schneider 2007). It is not clear, however, whether such an approach would still be feasible under the conditions of late modernity (Rampton 2006).

Where the movement of language resources is concerned, the influence of popular culture on language learning, in particular, needs to be given greater attention than has traditionally been the case in language contact studies. Stroud and Prinsloo's (2015: ix−x) observations are pertinent in this regard:

> Traditionally, studies of moving words have been filtered through bounded notions of language and community. From this perspective, linguistic forms are treated as belonging to (expert) speakers of the community, as displaying structural integrity and as being targets of acquisition by learners […]. But what if mobility, far from being an aberration or exception, is actually the normal state of affairs? What if instead of viewing linguistic mobility through the lens of a localist idea of language, the tables were turned and language was viewed through the lens of mobility? […] Power geometries of language are at work when linguistic forms travel − some travel well and others not so well. When the places and

> spaces across which words move are 'filled with codes, customs, rules, expectations and so forth' (Blommaert 2005: 73) with their indexical orders, the significance and value of particular forms of language are recalibrated.

The traditional focus in language contact studies has tended to be informed by the assumption that different language types (pidgin, creole, etc.) arise as bounded groups of speakers come to interact with one another. The dissemination of popular culture, in contrast, is not so much via interactions between groups of speakers as it is about how speakers come to appropriate culturally transmitted words and phrases as a result of textually mediated encounters.

The movements of people and language resources, then, raise interesting challenges and issues for our contemporary understanding of language contact. To better contextualize such challenges and issues, I first provide a brief review of key arguments that have been made in 'traditional' language contact studies. Here, I show how these rely on assumptions regarding the construction of distinct linguistic systems. As we have just noted, such assumptions are problematized in sociolinguistic studies informed by the notions of globalization and superdiversity. In the second half of the chapter, I discuss in further detail the implications of globalization and superdiversity for language contact. Phenomena such as globalization and superdiversity are significant because they highlight that, as a result of rapid change and increased social complexity, conventional ways of thinking about social organization no longer work, and this point also applies to how we think about language contact. In this regard, I point to recent sociolinguistic arguments that named languages have been denaturalized, that there is a need to approach language in terms of supervernaculars (Blommaert 2012; Blommaert and Rampton 2011), and that language itself has to be decentered.

2. Language contact studies and the denaturalization of language names and language systems

A key focus of language contact studies has been on establishing the "inner coherence of language 'systems'" (Matras 2009: 2). Theoretical disputes have tended to revolve around when and how a distinctly new language variety can be said to have come into existence. The discussions concerning the (disputed) properties of pidgins and creoles, as well as subsequent (controversial) attempts to establish the relevance of concepts such as a post-creole continuum (DeCamp 1971) and creoloid (Platt 1975), are all aimed at defining, in structural terms, the stages by which basic attempts at communication ultimately lead to the emergence of a full-fledged language.

This is perhaps easiest to see in the use of notions such as superstrate, substrate, and adstrate. Terms such as substrate and superstrate are oftentimes used to account for the ways in which different languages interact with one another, even though there is controversy over whether they are actually helpful in shedding light on the processes involved (Mufwene 1993, 2000). The superstrate is supposed to be the language upon which the pidgin or creole is based, and in colonizing situations is often taken to refer to the language of the colonizers because of its greater prestige. The superstrate is also assumed to be the lexifier language, that is, it contributes the basic vocabulary of the pidgin or

creole. In contrast, the substrate is supposed to be the language of the colonized people, and it may contribute, with varying degrees of prominence, to the lexical and grammatical structure of the pidgin or creole. Yet particularly when multiple languages are involved, it is not always clear whether distinct superstrate and substrate languages can be easily, or even coherently, identified. And of course, where the languages that are in contact cannot be said to be in any clear relationship of dominance, they are described as being adstrates.

The specific terms aside, it is undeniable that the invocation of these different types of '-strates' all depend on the presumption that there are preexisting, linguistically distinct, and autonomous systems that are brought into contact with one another. And the structural effects of the contact situation motivate the categorization of the distinct systems as specific types of '-strates'. However, as Matras (2009: 3) points out, "'[c]ontact' is, of course, a metaphor: language 'systems' do not genuinely touch or even influence one another. The relevant locus of contact is the language processing apparatus of the individual multilingual speaker and the employment of this apparatus in communicative interaction". Taking a somewhat different approach that places more emphasis on the interactional aspects of language contact, Arends, Muysken, and Smith (1995) propose that both pidgins and creoles come about because speakers are trying to simplify their language in order to be more easily understood by their interlocutors. This is known as the Foreigner Talk hypothesis, and it constitutes one clear departure from the otherwise strongly structural orientation that has characterized language contact studies. Yet even when the focus shifts to communicative interaction, the guiding question here is the extent to which the 'proper' grammatical rules and structures of a language have been simplified. That is, the full-fledged language still remains the point of reference, and divergences from its structures are accounted for in terms of the special demands of the contact situation. Implicit here, then, is the related question of how well speakers are able to successfully acquire the target language under contact situations. These kinds of questions make sense only if there are clear and stable affiliations between linguistic forms and a community. They become much more problematic, however, once mobility is taken to be the norm rather than the exception (Stroud and Prinsloo 2015, see above).

There still remains a strong tendency to gravitate towards talking about a variety as a reified entity, often to the point where speakers and linguistic forms are homogenized, and this tendency is identifiable even when the focus of discussion is precisely about language use in the context of global mobility. Thus, consider the English as a Lingua Franca (ELF) project (House 2010; Jenkins 2008; Seidlhofer 2007). As Park and Wee (2013: 342) point out, in one sense, to describe English as an international lingua franca is not at all contentious, since a lingua franca is simply the language that is used as a communicative bridge in situations where speakers come from different linguistic backgrounds, and English is undoubtedly used in multiple situations where it is not always the first language of the speakers involved. However, this is no longer the case with the ELF project (Park and Wee 2013: 343):

> The presumption of a varietal lingua franca is especially problematic in the international arena where it is difficult to posit an identifiable speech community, where 'there are too many different specific situations involved' and where 'each situation is also likely to have different individual speakers participating' from different cultural backgrounds (Park and Wee 2012: 51; see also references cited therein). It is this tendency to assume that it is coherent to speak of an ELF variety at the international level that has opened the ELF

> project to criticisms that its conception of ELF tends to be monolithic (Jenkins 2006: 161; Park and Wee 2012: 50–51; Prodromou 2008: 28).

The result is that we now have references to ASEAN ELF (Kirkpatrick 2007), South African ELF (Onraët 2011), East Asian ELF, European ELF, business ELF, and academic ELF (Jenkins 2008), even though the use of such labels, as Schneider (2011: 155) points out in his discussion of English in Asia, "unless understood very loosely and non-technically, implies more homogeneity than is warranted".

The issue is not just the labels, of course. An uncritical proliferation of language names indicates a dangerous tendency to assume that the names correspond to a distinct variety, and from there, the research program shifts into one of identifying the various structural properties associated with the named variety. The more fundamental question, and one that is especially important to keep in mind when looking at language contact in relation to globalization and superdiversity, is whether it is at all justifiable to assume that the named variety, and hence the distinct and autonomously existing linguistic system that it presumably refers to, actually exists (Bruthiaux 2003: 168).

Matras (2009: 144) provides a strong repudiation of this 'systems' assumption when he argues that the cognitive and individual perspective on language acquisition cannot be dissociated from the social environment, which provides the learning context:

> […] communicative interaction does not involve the activation or disengagement of 'systems', but rather constant choices among components of the full repertoire of linguistic structures (word-forms, constructions, phonetic realisations, and so on). In selecting among items, bilingual speakers are guided, naturally, by the meaning and function of those items, but also by their context-appropriateness. This involves assessing the immediate communicative interaction and classifying it as more or less strict in respect of expectations on code choice, as well as negotiating the integration of meanings and associations from other activity contexts, special effects and metalanguaging, authenticity, and so on. Control over the selection mechanism is a necessary pre-requisite for all these processes.

A child will learn to organize his/her linguistic repertoire (of words, phrases, and sounds) into different sets as a result of various social cues, such as addressee, setting, individual interlocutors, and purpose of interaction (Matras 2009: 42). Where there are consistent convergences between the different cues, there is a greater entrenchment of a particular linguistic repertoire (Matras 2009: 40).

A speaker's control over the ability to select appropriate aspects of his/her linguistic repertoire therefore develops in conjunction with various social and interactional cues. Such cues also guide the speaker in making metalinguistic decisions, including how to organize and differentiate his/her repertoire into distinct, named varieties.

3. On fluid multilingualism and urban vernaculars

The problem with assuming the existence of linguistic systems is further reinforced in Pennycook's (2003) critique of the World Englishes enterprise, which he faults for not recognizing the performative dimensions of using English. In his discussion of 'raplish', Pennycook (2003: 529) argues that the use of English in rap music among Korean or

Japanese rappers, for example, is evidence that speakers utilize a range of linguistic resources (Korean, Japanese, and African American Vernacular English) to signal affiliations that operate simultaneously at the global, regional, national, and local levels rather than just at national boundaries.

This skepticism about distinct systems is further driven home in more recent work arising from investigations of language use in the workplace, where speakers mix resources from Japanese and English. As Otsuji and Pennycook (2010: 241) observe, "[a]t the very least, then, we can note that such instances of English/Japanese mixed code use derive not so much from the use of different first and second languages but rather as the result of a mixed Japanese/English code becoming the lingua franca of the workplace". Otsuji and Pennycook (2010: 241) also highlight how this questioning of distinct systems and boundaries has been gaining momentum in sociolinguistics and applied linguistics: "A recent movement in bi- and multilingual studies has been to shift away from a focus on how distinct codes are switched or mixed, in favour of an interest in how boundaries and distinctions are the results of particular language ideologies and how language users manipulate the multilingual resources they have available to them". Indeed, there has been a call in both sociolinguistics and applied linguistics for a better appreciation of the "fluid multilingual realities of today's world" (Weber and Horner 2012: 117), a fluidity that has resulted in a minor proliferation of terms, having been variously characterized as translanguaging (García 2009), polylanguaging (Jørgensen 2008), metrolingualism (Otsuji and Pennycook 2010), and translingualism (Canagarajah 2013). Pennycook (2016) describes this as the 'trans-super-poly-metro movement'. While there are minor differences in how the individual terms are employed (e.g. 'translanguaging' has a strong pedagogical orientation, while 'polylingualism' stresses the need to attend to individual linguistic features rather than whole languages), the commonalities are more predominant, and hence, the collective sense of a 'movement'. There is a collective emphasis on how speakers appear to mix the resources associated with conventionally established linguistic varieties, to such an extent that any presumption of an existing autonomous linguistic system must be met with skepticism.

For example, Canagarajah (2015: 41) points to the "hybrid form of literacy activity" found on a Facebook page created by students from a South African township school, where "the participants use the conventions of text messaging (featuring abbreviations and icons) in addition to mixing English and Xhosa". Jacquemet (2005: 269) describes the communicative practices at a family reunion, where "the Albanian language blended with English, Italian, and occasionally German […]. The TV was on, broadcasting in Spanish, while from another room you could hear the beat of Nuyorican hip hop […]", and because some of the family members were keen to learn Italian, these individuals "in their peer group […] displayed these newly acquired linguistic skills in a mixed idiom of Albanian, Italian, English, and personal slang" (Jacquemet 2005: 270). Finally, Otsuji and Pennycook (2010: 241) highlight the mixed use of Japanese and English in an Australian workplace where none of the interlocutors are Japanese, and note that "such instances of English/Japanese mixed code use derive not so much from the use of different first and second languages but rather as the result of a mixed Japanese/English code becoming the lingua franca of the workplace".

In a strong endorsement of the need to give greater attention to the theoretical ramifications of such fluidity in multilingual practices, Blommaert (2012: 13) states that:

> Old and established terms such as 'codeswitching', and indeed even 'multilingualism', appear to rapidly exhaust the limits of their descriptive and explanatory power in the face of such highly complex 'blends' (cf. Sharma & Rampton 2011; Backus 2012; Creese & Blackledge 2010). And not only that: the question where the 'stuff' that goes into the blend comes from, how it has been acquired, and what kind of 'competence' it represents, is equally difficult to answer. Contemporary repertoires are tremendously complex, dynamic and unstable, and *not* predicated on the forms of knowledge-of-language one customarily assumes, since Chomsky, with regard to language (Blommaert & Backus 2012).

It is actually useful in this regard to consider what it means to speak of English as the global language *par excellence.* As Blommaert (2013: 8) points out, mobility creates unpredictability, because language gets dislodged from its traditional settings, and its traditional functions become distorted by the processes of mobility as language gets inserted into new settings (see also Blommaert 2010: 197). This process of insertion changes the language itself. Blommaert (2012) suggests that the spread of language, especially when this process occurs via the Internet and various mobile communication devices, requires the acknowledgement of a new kind of sociolinguistic object, that of the supervernacular. The supervernacular is needed to make sense of language use among communities that are deterritorialized and dynamic, and where membership may be elastic because such communities lack many of the characteristics associated with more traditional communities, such as fixed territorial boundaries, physical interactions, and shared cultural backgrounds and values (Blommaert 2012: 3). Blommaert (2012: 5) explicitly wants to subsume under the notion of a supervernacular the kinds of fluid language uses identified by Jørgensen (2008), Otsuji and Pennycook (2010), and Rampton (2010) as polylanguaging, metrolingualism, and urban vernaculars, respectively.

Blommaert (2012: 4) emphasizes that the 'super' in 'supervernaculars' is to be understood differently than the 'super' in 'superdiversity':

> Observe, in passing, that the 'super' in 'supergroups' and 'supervernaculars' if [sic] of a different order than that in 'superdiversity'. While 'super' in the latter stands for 'hyper', i.e. a degree higher than that encountered before or elsewhere, the 'super' in 'supergroups' and 'supervernacular' is an equivalent of 'trans-' and refers to communities and semiotic complexes whose composition and circulation transcend those of other semiotic complexes. The important point is to realize that supervernaculars only occur as dialects. Hence the 'vernacular' in the term. *There is no 'real' supervernacular, other than the dynamic complex of emerging, stabilizing and changing dialects we actually observe, hear, speak, read and write* [emphasis mine].

The final sentence in the above emphasizes that the supervernacular is not a reified linguistic object, much less an autonomous and pre-existing system. It is intended to capture the highly dynamic, and to varying degrees stabilized sets of conventions that speakers negotiate, produce, and contest as they move across multiple community memberships. As an example of a supervernacular, Blommaert (2012: 8) discusses mobile texting codes, whose conventions have been influenced by technological, ergonomic, and economic factors. For example, the use of *lkn* ('looking'), *CU* ('see you'), and various emoticons are all by now conventionalized semiotic resources that constitute the supervernacular of mobile texting.

The introduction of such terms is not without criticism, however. Edwards (2012: 38) describes these as "barren verbiage [lacking in] novelty". Pavlenko (2018) questions if

the focus on mobility of languages and of people really represents a significantly new phenomenon, or whether too much is being made of something that has already been going on for a long time in human history. And even more sympathetic commentators, such as Blackledge and Creese (2014), have suggested that the emphasis on the unbounded nature of language use might be traceable to Mikhail Bakhtin's notion of heteroglossia. However, aside from the issue of whether we are looking at anything conceptually new, there is also the question of just how widespread or constrained the language practices that these terms describe actually are.

Many of the language practices that these terms describe are observed in spaces where such mixing is more likely to be tolerated or encouraged, as in pedagogical arenas concerned with creative writing, informal workplace banter or family gatherings (see also Pennycook 2016: 205). But as Wee (2016: 339) points out, many organizations tend to construe multilingualism in fixed rather than fluid terms, which means that learners still need to be "prepared for institutional realities where the mixing of resources associated with distinct language systems is frowned upon and even penalized". Busch (2012: 507) makes a similar observation when she points out that many of these terms rely, implicitly or otherwise, on Gumperz's notion of a repertoire and goes on to caution that their emphases on individual linguistic agency run the risk of ignoring the continued normative effects of more traditional conceptions of language as bounded entities: "Translanguaging as a linguistic practice playfully combines elements that represent references to different linguistic and social categorizations or national stereotypes. By overstepping categories of a normative character, it does not eliminate these but reinvokes them, as it were. Moreover, opposed to the practices of translanguaging are institutional practices that aim at monolingualization and homogenization and determine inclusions and exclusions". In other words, however mobile speakers and languages may be, these still need to be physically located somewhere. Depending on who these individuals are and where they are located (or attempting to be located), there will always be localized constraints that they need to negotiate − which therefore means that any discussion about agency will need to be carefully qualified.

Therefore, while there is undoubtedly a need to seriously consider the challenges posed by mobility, fluidity, and the breaking of boundaries, both at the levels of community membership and of communicative practices, this has to be balanced against the fact that more traditional and orthodox assumptions about languages are still highly influential. Such orthodoxies − languages are bounded, autonomous systems; these bounded systems are themselves linked in essential ways to community identities and boundaries; and transgressions across language boundaries are viewed as a lack of competence − are not likely to disappear in the near future.

4. Conclusion

Blommaert (2005: 11) observes approvingly that "[i]t is one of sociolinguistics' great accomplishments to have replaced a uniform and homogeneous notion of 'language' − 'English', 'French' etc. − by a fragmented one, and to have explained why this fragmentation is necessary". While I share this approval, it may be somewhat premature to speak of this as an 'accomplishment' rather than a work in progress, since modernist

conceptualizations that insist on the reification of language systems still remain influential.

We have to consistently understand language as the product of human activity, and changes to language − including what we may each differentially believe to be 'language' − as the results of changing human activities. Language does not preexist these activities, but emerges from them. The lesson to be gleaned from globalization and superdiversity is that our language is far from 'complete', and more fundamentally, that to think in terms of completeness is to misconstrue the nature of language. Yet even if the dramatic changes in the nature of social and linguistic diversity pose serious problems for the desire to see languages as neatly compartmentalized systems, the inability to contemplate or accept alternative understandings of language means that this modernist conceptualization is unlikely to disappear anytime soon.

5. References

Arends, Jacques, Pieter Muysken & Norval Smith
1995 *Pidgins and creoles: An introduction*. Amsterdam: John Benjamins.

Blackledge, Adrian & Angela Creese
2014 Heteroglossia as practice and pedagogy. In Adrian Blackledge & Angela Creese (eds.), *Heteroglossia as practice and pedagogy*, 1−20. Dordrecht: Springer.

Blommaert, Jan
2005 *Discourse: A critical introduction*. Cambridge: Cambridge University Press.

Blommaert, Jan
2010 *The sociolinguistics of globalization*. Cambridge: Cambridge University Press.

Blommaert, Jan
2012 Supervernaculars and their dialects. *Dutch Journal of Applied Linguistics* 1(1). 1−14.

Blommaert, Jan
2013 *Ethnography, superdiversity and linguistic landscapes: Chronicles of Complexity*. Clevedon: Multilingual Matters.

Blommaert, Jan & Ben Rampton
2011 Language and superdiversity. *Diversities* 13(2). 1−21.

Blommaert, Jan & Ben Rampton
2016 Language and superdiversity. In Karel Arnaut, Jan Blommaert, Ben Rampton & Massimiliano Spotti (eds.), *Language and superdiversity*, 21−48. New York: Routledge.

Bruthiaux, Paul
2003 Squaring the circles: Issues in modelling English worldwide. *International Journal of Applied Linguistics* 13(2). 159−78.

Busch, Brigitta
2012 The linguistic repertoire revisited. *Applied Linguistics* 33(5). 503−523.

Canagarajah, A. Suresh
2013 *Literacy as translingual practice: Between communities and classrooms*. New York & London: Routledge.

Canagarajah, A. Suresh
2015 Negotiating mobile codes and literacies at the contact zone. In Christopher Stroud & Mastin Prinsloo (eds.), *Language, literacy and diversity*, 34−54. London: Routledge.

DeCamp, David
1971 Towards a generative analysis of a post-creole speech continuum. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 349−370. Cambridge: Cambridge University Press.

Edwards, John
2012 *Multilingualism: Understanding linguistic diversity.* London: Continuum.
Faist, Thomas
2000 *The volume and dynamics of international migration and transnational social spaces*. Oxford: Oxford Scholarship Online.
García, Ofelia
2009 Education, multilingualism and translanguaging in the 21st century. In Ajit Mohanty, Minati Panda, Robert Phillipson & Tove Skutnabb-Kangas (eds.), *Multilingual education for social justice: Globalising the local*, 128–145. New Delhi: Orient Blackswan.
Giddens, Anthony
1990 *The consequences of modernity.* Cambridge: Polity.
Giddens, Anthony
2002 *Runaway world: How globalization is reshaping our lives*, 2nd edn. London: Profile Books.
House, Juliane
2010 The pragmatics of English as a lingua franca. In Anna Trosborg (ed.), *Pragmatics across languages and cultures* (Handbook of Pragmatics, vol. 7), 363–387. Berlin: Mouton de Gruyter.
Jacquemet, Marco
2005 Transidiomatic practices: Language and power in the age of globalization. *Language & Communication* 25(3). 257–277.
Jenkins, Jennifer
2006 Current perspectives on teaching world Englishes and English as a Lingua Franca. *TESOL Quarterly* 40(1). 157–181.
Jenkins, Jennifer
2008 English as a lingua franca. Keynote presented at the JACET 47th Annual Convention, Waseda University, 11–13 September.
Jørgensen, Jens Normann
2008 Poly-lingual languaging around and among children and adolescents. *International Journal of Multlingualism* 5(3). 161–176.
Kennedy, Paul
2001 Introduction: Globalization and the crisis of identities? In Paul Kennedy & Catherine J. Danks (eds.), *Globalization and national identities: Crisis or opportunity?* 1–28. New York: Palgrave.
Kirkpatrick, Andy
2007 *World Englishes: Implications for international communication and English language teaching*. Cambridge: Cambridge University Press.
Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.
Mufwene, Salikoko S.
1993 *Africanisms in Afro-American language varieties*. Athens: University of Georgia Press.
Mufwene, Salikoko S.
2000 Creolization is a social, not a structural process. In Ingrid Neumann-Holzschuh & Edgar Schneider (eds.), *Degrees of restructuring in creole languages*, 65–84. Amsterdam: John Benjamins.
Mufwene, Salikoko S.
2001 *The ecology of language evolution*. Cambridge: Cambridge University Press.
Onraët, Lauren Alexandra
2011 *English as a lingua franca and English in South Africa: Distinctions and overlap*. Stellenbosch: University of Stellenbosch. MA thesis.

Otsuji, Emi & Alastair Pennycook
2010 Metrolingualism: Fixity, fluidity and language in flux. *International Journal of Multilingualism* 7(3). 240–254.
Park, Joseph Sung-Yul & Lionel Wee
2012 *Markets of English: Linguistic capital and language policy in a globalizing world.* New York: Routledge.
Park, Joseph Sung-Yul & Lionel Wee
2013 Linguistic baptism and the disintegration of ELF. *Applied Linguistics Review* 4(2). 339–59.
Pavlenko, Aneta
2018 Superdiversity and why it isn't: Reflections on terminological innovation and academic branding. In Stephan Breidbach, Lutz Küster & Barbara Schmenk (eds.), *Sloganizations in language education discourse*, 142–168. Bristol: Multilingual Matters.
Pennycook, Alastair
2003 Global Englishes, rip slyme, and performativity. *Journal of Sociolinguistics* 7(4). 513–533.
Pennycook, Alastair
2016 Mobile times, mobile terms: The trans-super-poly-metro movement. In Nik Coupland (ed.), *Sociolinguistics: Theoretical debates*, 201–216. Cambridge: Cambridge University Press.
Perrons, Diane
2004 *Globalization and social change: People and places in a divided world.* London/New York: Routledge.
Platt, John
1975 The Singapore English speech continuum and its basilect 'Singlish' as a 'creoloid'. *Anthropological Linguistics* 17(7). 363–74.
Prodromou, Luke
2008 *English as a lingua franca*. London/New York: Continuum.
Rampton, Ben
2006 *Language in Late Modernity.* Cambridge: Cambridge University Press.
Rampton, Ben
2010 From 'multi-ethnic urban heteroglossia' to 'contemporary urban vernaculars'. *Working Papers in Urban Language and Literacy* Paper 61. London: King's College (Centre for Language, Discourse and Communication).
Schneider, Edgar
2007 *Postcolonial English*. Cambridge: Cambridge University Press.
Schneider, Edgar
2011 English into Asia: From Singaporean ubiquity to Chinese learners' features. In Michael Adams & Anne Curzan (eds.), *Contours of English and English language studies*, 135–156. Ann Arbor: University of Michigan Press.
Seidlhofer, Barbara
2007 Common property: English as a lingua franca in Europe. In Jim Cummins & Chris Davison (eds.), *The international handbook of English language teaching*, 137–153. New York: Springer.
Stroud, Christopher & Mastin Prinsloo
2015 Preface. In Christopher Stroud & Mastin Prinsloo (eds.), *Language, literacy and diversity: Moving words*, ix–xiii. London: Routledge.
Vertovec, Steven
2007 Super-diversity and its implications. *Ethnic and Racial Studies* 30(6). 1024–1054.
Vertovec, Steven
2010 Towards post-multiculturalism? Changing communities, contexts and conditions of diversity. *International Social Science Journal* 61(199). 83–95.

Wade, Robert Hunter
2001 Is globalization making world income distribution more equal? *London School of Economics DESTIN Working Paper* 01–01. London: LSE.
Weber, Jean-Jacques & Kristine Horner
2012 *Introducing multilingualism: A social approach.* New York/London: Routledge.
Wee, Lionel
2016 Are there zombies in language policy? Theoretical interventions and the continued vitality of (apparently) defunct concepts. In Nik Coupland (ed.), *Sociolinguistics: Theoretical debates*, 331–48. Cambridge: Cambridge University Press.

Lionel Wee, Singapore (Republic of Singapore)

29. States, nations, and language contact

1. Introduction
2. Linguistic diversity in premodern empires
3. The social basis of linguistic nationalism
4. Administering nationalizing states
5. Linguistic nationalism and corpus planning in the wake of European empires
6. Prospects for the future
7. References

1. Introduction

State formation and nationalism affect language contact in diverse ways, since different political ideologies, articulated within different state structures, generate different types of linguistic politics. Nevertheless, a few key issues dominate. First and foremost, linguistic nationalism generally concerns not spoken language, but written language. The literacy rate is therefore important: language contact between illiterate populations creates different political conflict than does language contact between literate populations. State sponsorship is also important: all states are multilingual, but to different degrees, confronting governments with different administrative challenges. Which literary standardization or standardizations does the government use in educational and/or state administration? To what extent does the government seek to impose its favored language(s) on speakers of other languages? How do minority language communities respond to relative disenfranchisement? The answers to such questions vary from place to place, but also from one historical era to the next. The strategies adopted to resolve such difficulties illustrate the varied interactions between linguistic and national ideologies.

Since state structures determine the politics of language contact, this chapter has thematic sections, each focusing on different political circumstances. It begins with an overview of premodern empires, from ancient times to European colonialism (Section 2). Section 3 considers how a series of social transformations, variously characterized as ‘industrialization’ or ‘modernization’, generated linguistic nationalism, and particularly the ‘language-dialect’ disputes that so often characterize modern linguistic politics. Section 4 considers the role of linguistic nationalism in declining multiethnic empires, and

https://doi.org/10.1515/9783110435351-029

the national states that subsequently emerged from them. Section 5 examines linguistic politics in postcolonial states. Section 6 tentatively speculates about the future of linguistic politics in the current era of globalization.

2. Linguistic diversity in premodern empires

The importance of standardized languages has grown steadily through history, but even the earliest states of ancient Mesopotamia required some means of record keeping and written communication (Radner 2014). When the state consisted primarily of a royal court, a temple, and an army, literacy was an exceptional ability. While premodern states often encompassed significant linguistic diversity, their rulers were typically comfortable governing multilingual populations, since only a small number of scribes and administrators needed to know the state language. Indeed, literacy was so esoteric a skill that a scribal caste might outlive the language it mastered. Sumerian, the world's first written language, endured as a medium of record keeping and culture long after its decline as a spoken language (Michalowski 2008: 8).

Many premodern empires adopted the literary language of a conquering people. The polyglot Empire of Alexander the Great, and his Empire's polyglot successor states, used Greek. The prestige of conquerors and rulers extended to their language. In Ptolemaic Egypt, the appeal of Greek civilization, art, and philosophy combined with practical considerations to spread the Greek language widely: Egyptians needed to speak Greek to access the court system. Overall, "Greek influence was profound" (Bauschatz 2013: 331). Linguistic influence struck particularly deep roots if the language of conquerors became associated with a shared religion. Indeed, religious prestige unaccompanied by military conquest spread Sanskrit throughout southeast Asia (Pollock 2006: 115–134). Both imperial conquest and religious conversion contributed to the spread of Latin in Western Europe (Millar 2010: 71–84, 89–98, 114–118), the spread of Arabic in the Middle East and North Africa (Versteegh 2014: 126–171), and the spread of Spanish in the Americas (Brown 1999).

Premodern conquest often led to language mixing, multilingualism, and language shift. Several factors determined the speed and extent of assimilation, but relative population size was important: a densely populated region more readily assimilates its conquerors. Assimilation also spreads more rapidly among related varieties. Indo-European Persians or Punjabis less easily adopted Arabic than did speakers of Semitic languages like Aramaic and Syriac, even if Syrian Arabic underwent a certain creolization (Versteegh 1984: 73–76).

Several premodern empires used state languages not widely spoken by their subject populations. The administrative language of the Ottoman state, while fundamentally a variety of Turkic, borrowed so many words and expressions from Arabic and Persian, prestigious languages of Islamic erudition, that ordinary Turks required special training to master it (Doğançay-Aktuna 2004). In Christian Europe, Latin persisted as an official state language long after its decline as a spoken language. In Hungary, and particularly in Croatia, a predominantly Slavic province of the Habsburg monarchy, Latin was retained until the 19th century (Šubarić 2015). In the multiethnic, multiconfessional, and multilingual Ottoman and Habsburg Empires, all aspiring bureaucrats learned the state

language as a second language. From the government's perspective, linguistic continuity facilitated administrative continuity. The expense of the necessary schooling helped social elites maintain their monopoly on lucrative and prestigious posts.

The history of Chinese most dramatically illustrates how a prestigious and well-established state language can survive centuries after spoken varieties diverge from an archaic written standard. Han China encompasses several different linguistic regions, each associated with what in modern Chinese is called a *fangyan*. The term *fangyan* originally referred to languages other than Chinese (Yuming 2015: 158), but has since come to denote Chinese varieties whose mutual intelligibility cannot be assumed (DeFrancis 1984: 57–58). Regional variants of literary Chinese, however, differ much less than variants of spoken Chinese. As early as 213 BCE., the Qin Emperor Shihuang, in perhaps the world's earliest recorded case of language planning, proscribed a specific literary style and ordered the burning of books which did not conform to it (Hannas 1997: 12–15; Norman 1988: 1–22). Though the Qin Emperor is remembered as a tyrant, his linguistic reforms endured. While the Chinese literary tradition contains diversity, it has retained a surprising level of homogeneity through centuries of history.

The extraordinary unity and longevity of literary Chinese derives partly from the logographic nature of the Chinese character script. Written Chinese originally developed from pictograms, and expanded on the basis of the rebus principle (Boltz 1994: 168–177; DeFrancis 1984: 74–82, 136–140). The characters have since lost much of their phonetic nature, and indeed are sufficiently distant from modern speech that speakers of all *fangyan* must struggle to attain mastery, much as speakers of Italian, French, and Spanish would struggle to master Latin.

The archaic and ideographic Chinese character script has persisted as a literary and administrative language in the age of nationalism partly because replacing it with an easier-to-learn phonetic alphabet would require the government to select a standard pronunciation from one *fangyan*, inevitably associated with one particular region, thus disenfranchising speakers from other parts of the country. After the triumph of the Communist party in 1949, Mao Zedong's government briefly considered replacing the character script with Pinyin, an alphabetic script using the Latin alphabet. Phonetically, Pinyin was based on the pronunciation of Beijing, and northern Chinese found it easier than the traditional character script. Speakers of southern Chinese *fangyan*, however, found it difficult and resisted (Peterson 2011: 107–17). Fearing separatism, the government backed down and concentrated its reformist impulses on simplifying the characters.

3. The social basis of linguistic nationalism

New forms of linguistic conflict appeared in the modern era. As modern states expanded into new social and economic roles, they increasingly replaced prestigious traditional languages with easy-to-learn living languages. Economic growth, evolving communications technology, and rising literacy rates spawned postal networks, state banks, customs offices, school systems, libraries, museums, pension schemes, and countless other state ministries which collectively required an army of literate clerks. As the demand for literate state officials grew, impoverished subalterns realized that literacy would bring their children social mobility and financial security. As the children of illiterate parents

became educated, they contributed to the rise of nationalism. Weber (1976: 303–338) has vividly described how the spread of education helped transform "peasants into Frenchmen"; the spread of literacy has similarly transformed peasant populations all across the world.

The sudden importance of schooling generated unexpected conflicts in polyglot empires that had previously found multilingualism unproblematic. If the conscious cultivation of national linguistic distinctiveness was initially a "scholarly interest" pursued by small numbers of savants, it often inspired "patriotic agitation", or even a "mass national movement", to use the terminology Hroch (1985: 23) introduced in his influential study of how national movements emerge and develop. Phase-based models of so-called 'national revival' summarize a complicated process, whose linguistic counterpart, sometimes described as a 'language revival', consists mostly of standardization, codification, and corpus planning (Bentahila and Davies 1993; Paulston, Chen, and Connerty 1993; see also Chapters 30 and 32).

Language revival, like national revival, has also been analyzed with phase theories. Haugen (1966: 933, 1987: 59–70) influentially posited a fourfold model of language standardization and codification, according to which activists first chose a norm to standardize, then codify it through corpus planning, then implement the standard by using it (e.g. in newspapers, textbooks, and so forth), and finally elaborate the standard by expanding its functions (e.g. by devising new technical vocabularies). Haugen's model has in turn inspired further theorists, who have added new phases or otherwise reformulated Haugen's basic model (Muljačić 1993; Neustupný 1970). Hroch (1994: 12–17), the aforementioned nationalism theorist, also developed a six-stage model.

Both national revival and language revival arise from economic and social change, but both also have an ideological dimension. While medieval scholars had prided themselves on their mastery of dead languages, such as Latin, Arabic, or Sanskrit, modern patriots began professing their love of their native language, even before that language underwent formal standardization and codification. Promoting the mother tongue became a national virtue. The influential work of German theologian Johann Gottfried von Herder held that every nation in the world played some unique role in God's unfolding plan, and expressed its unique character through its unique language (Herder 1784–1791). Though Herder wrote primarily for German readers, his ideas found a broad audience elsewhere (Andraschke and Loos 2002). Intellectuals from minority language communities wrote grammars and dictionaries, collected folk tales, composed poetry, and more generally developed ideologies 'in praise of the beloved language', to use the phrase of Fishman (1997).

During the early phases of language revival, patriot linguists often disagreed about how exactly to standardize. Savants have variously disagreed about which alphabet or orthography to adopt, whether to use an etymological or phonetic spelling system, proper standard pronunciation, and a host of other similar questions. Some revivals have floundered on technical linguistic issues: the Cornish language revival, for example, has suffered because the tiny Cornish intelligentsia cannot agree on which standardization to adopt (Deacon 2006).

Technical disagreements may overlap with different notions about the extent of the national language. A particularly contentious issue arises when speakers of related varieties must decide whether or not they share a language, or speak different albeit related languages. Linguistic diversity is ubiquitous, but both its degree and its perceived impor-

tance vary. Even non-activist scholars may disagree in good faith about the correct classification within a language family. As Lieber (1997: 56) noted, "in almost any field of classification there are lumpers and splitters". Querulous activists, however, may treat linguistic debates as a proxy for various political agendas. Such disagreements strike to the heart of national aspirations. Linguistic differences are sometimes downplayed, and are other times emphasized as markers of group distinctiveness. Metropolitan intellectuals may view regional varieties as mere 'dialects' of a national language, while intellectuals from peripheral or disadvantaged regions may prefer to imagine themselves as speakers of a distinct and persecuted 'language'. Just as disagreements arise about whether a given province deserves political autonomy or independence, so too do they arise about where a linguistic frontier lies, or whether it exists at all.

Conflicts over the status of 'languages' and/or 'dialects' have certain generalizable features. Advocates of both linguistic separatism and linguistic unity typically proclaim their preferred linguistic taxonomy an objective reality, dismissing their opponents as seeking to 'invent' an 'artificial language' from 'political' motives. Such accusations are best understood as strategies of delegitimization. National revival and language revival, with their concomitant claims to 'language-hood', or to a 'national language', are inherently political. Indeed, since all national languages require standardization and corpus planning, Hobsbawm (1992: 54) rightly noted that they are "almost always semi-artificial constructs and occasionally, like modern Hebrew, virtually invented".

Over time, consensus opinion about language-hood can change dramatically, both among patriot activists and among target populations. Varieties once imagined as mere dialects may win recognition as proper 'languages'. Kloss (1967) memorably theorized this process as the creation of *Ausbau* languages, or creating 'languages by development'. Kloss distinguished *Ausbau* languages, which differ because they have undergone language planning and codification, from *Abstand* languages, or 'languages by distance', whose structure and vocabulary are intrinsically different from each other. Spanish and Basque, for example, would be *Abstand* languages, while Spanish and Catalan would be *Ausbau* languages. Kloss, unfortunately, provided no catchy terms for discussing uncodified varieties before any codification processes.

New *Ausbau* languages are continually emerging in various parts of the world. South Slavic provides particularly vivid examples. During the 19th century, belief in the linguistic unity of South Slavic varieties inspired considerable national work: the so-called 'Illyrian movement' posited a single language for all South Slavs (Despalatović 1975). Bulgarian and Slovenian became firmly established as separate languages in the 19th century, and Macedonian after the Second World War. For most of the 20th century activists and savants posited a single 'Serbo-Croat' language. Enduring religious divides in the multiconfessional Balkan Peninsula, combined with military conflicts at the end of the 20th century, eventually fragmented the 'Serbo-Croat' intelligentsia as well. Successor intelligentsias now espouse not only distinct Serbian and Croatian 'languages' (Greenberg 2004; Okey 2004), but distinct Bosnian and Montenegrin 'languages' as well (Nakazawa 2015; Troch 2014).

Ideologies of the national language best take root when they articulate collective interests. Speakers of a minority language, and particularly intellectuals who have mastered a recently codified literary standard, stand to benefit if state institutions grant that codified standard some official status. The most outspoken advocates of a language revival are typically those scholars and intellectuals with the most to gain. While Hobs-

bawm (1992: 96, 117–18) warned scholars not to "reduce linguistic nationalism to a question of jobs", he rightly emphasized that a government's decision to support one or another written language in "office and school" can profoundly affect people's material interests.

Even when a consensus exists that two related varieties form separate 'languages', the border between them may change. When the region of Scania belonged to Denmark, its inhabitants saw themselves as speakers of Danish. After the Swedish conquest, Scanians slowly reclassified their speech as dialects of Swedish (Ohlsson 2005: 1366–1367). Elites from a linguistic region contested by rival power centers may develop a sense of particularism. Both the Serbian and Bulgarian national movements claimed Slavic Macedonia, but Macedonian intellectuals eventually declared themselves neither Serbian nor Bulgarian but rather distinctively Macedonian (Tomić 1992).

4. Administering nationalizing states

Language politics in the age of nationalism therefore developed unique features. Firstly, linguistic nationalism challenged longstanding administrative practices in multilingual states. Agitators from minority linguistic communities, typically presenting themselves as representatives of a nation, attempted to win for their language an official status in state institutions: schools, court systems, local administration, and so on. They thus demanded administrative reforms, and typically an autonomous district in which their language would enjoy official status. It is important to recognize that demands for administration in a minority language need not imply demands for political independence: advocates of language revival can achieve many of their goals with an autonomous administrative district within a larger multiethnic state.

Demands for linguistic autonomy within a distinct administrative unit can be both made and resisted in the name of national unity. When, for example, Emperor Joseph II abandoned Latin as the administrative language of the Habsburg monarchy, German intellectuals stood to benefit. Hungarian noblemen, however, feared the change would disadvantage them in the competition for government posts. Representatives of the Hungarian nobility thus patriotically demanded Hungarian become the official language within the Kingdom of Hungary, the largest of the Monarchy's many constituent territories. When the polyglot Kingdom of Hungary adopted Hungarian as an official language, however, leading figures from the kingdom's Croatian, Romanian, and Slovak minorities resisted linguistic disenfranchisement by demanding their own autonomous administrative regions on national grounds (Kamusella 2009: 434–39, 447–453). When Croatia subsequently received autonomy within Hungary, the Croatian parliament introduced Croatian into local government, but Serbs living in Croatia then invoked national rhetoric to demand further autonomy within Croatia (Miller 1995).

Many successful multilingual states acknowledge different official languages in different provinces. Indeed, federal units, each with a single predominant language, characterize the canonical examples of successful multilingual states. Canada grants official dominance to French within Quebec (Cardinal 2015). Belgium is divided into four areas with distinctive language regimes. There is a large territory using Flemish (Dutch), a somewhat smaller territory using French, and a tiny region officially using German.

Brussels, the capital city, is officially multilingual (van Goethem 2010: 67–84, 153–166, 199–201; Peersman, Rutten, and Vosters 2015; Sonntag 1991). Switzerland has a complex federal system, yet while all its various cantons reward official multilingualism, linguistically disenfranchised minorities still generate political tension. The hegemony of the German language in canton Berne, for example, contributed to a francophone separatist movement, leading to the establishment of canton Jura in 1979 (McRae 1983: 185–212). Safran (2010: 58) rightly observed that "bilingualism is not a zero-sum game", yet nationalized struggles to introduce a given language into school classrooms, courtrooms, or government offices still have clear winners and losers. A person is more easily bilingual than a classroom assignment, a legal opinion, or a ministerial filing system.

Should community leaders from a linguistic minority achieve an independent state, or have it thrust upon them, the status of and institutional support given to different languages may change suddenly and dramatically. The collapse of multinational empires usually exacerbates linguistic conflicts. During the 20th century, new states have increasingly claimed to be nation-states, and political leaders in nation-states usually feel more pressure to pander to dominant linguistic communities at the expense of linguistic minorities. At the end of the First World War, several newly formed states introduced the language of the state's titular nation. Interwar Poland, for example, imposed the Polish language on significant minorities of Germans, Ukrainians, Belarussians, and Yiddish-speaking Jews (Kamusella 2009: 598–609). Members of minority communities responded with grievance, leading to political instability. Similar vicious cycles arose in other post-imperial states.

Several of the newly independent states found that the administrative requirements of statehood posed unexpected linguistic demands. Corpus planning and/or language purism resulted. The Turkish republic founded by Mustafa Kemal, widely known as Atatürk, provides a particularly dramatic example: Kemalist reformers replaced Arabic and Persian loanwords with Turkish neologisms, and abandoned the Arabic alphabet for a 'New Turkish' alphabet based on Latin letters (Çolak 2004). Peti-Stantić (2009: 77–78, 82) uses the term "restandardization" to describe such "sudden changes at one or more levels of the standard language". Kemalist patriots preoccupied with restandardization showed little patience toward linguistic minorities such as the Kurds, and implemented radical assimilationist policies. Over time, however, attempts to suppress the Kurdish language have provoked rather than suppressed Kurdish nationalism (Hassanpour 1992).

Increased standardization sometimes generated new conflicts between speakers of similar varieties. In interwar Czechoslovakia, the state posited a 'Czechoslovak language' with two standard written variants, Czech and Slovak. In Yugoslavia, the two written variants of an imagined 'Serbo-Croat language' even used different alphabets. In practice, however, people whose schooling took place in different variants competed in different job markets, read different newspapers, and generally adopted different national orientations. An imagined 'national language' may encompass a great variety of spoken language. Ideologies of linguistic unity rarely survive the competition for scarce resources which inevitably results when different social constituencies use different written standardizations. Such conflicts illustrate the different political meanings of spoken and written language, confirming that linguistic nationalism contests written literary standards more vigorously than regional spoken accents.

5. Linguistic nationalism and corpus planning in the wake of European empires

European colonial empires imposed European languages on indigenous peoples all around the world, but the collapse of European imperialism forced patriots from formerly colonized peoples to consider the linguistic needs of newly independent states. Several factors influenced postcolonial language policies, but varying levels of corpus planning presented policy makers with different options.

In the Americas, independence leaders tended to be native speakers of the colonial language, and typically descended from European colonists. George Washington and Simón Bolívar, whatever their motives, did not take up arms against their respective colonial powers because of linguistic grievances. Only after attaining independence did American republics develop linguistic nationalism. While American patriotism inspired Noah Webster's language planning, his famous American dictionary appeared in 1828, decades after the country's declaration of independence (Stanford 2016: 89–91). Distinctively Brazilian dictionaries did not appear until the 20th century (Baxter 1992: 19–22).

Speakers of indigenous American languages have often attempted language revival. Scholarly interest, however, has concentrated primarily on preservation, rather than corpus planning (Frawley, Hill, and Munro 2002). The Navajo language, for example, is taught in schools, but the Navajo intelligentsia writes more in English than in Navajo (House 2002). Access to government resources also remains limited: teachers of Nahautl, the largest indigenous language in Mexico, offer workshops mostly in their private time (Messing and Nava 2016: 86). Indigenous American languages are rarely used in government offices. Most are endangered.

Though indigenous languages are more widely spoken in Africa, most postcolonial states have kept colonial languages as official languages. Colonial boundaries in Africa rarely coincided with linguistic frontiers. Independent African states typically contained many different language groups. Promoting any one language threatened to spark ethnolinguistic conflict (Laitin 1992). The former colonial languages are neutral with regard to local power struggles, enjoy associations with power and 'modernity', and enable advanced speakers to access prestigious international opportunities. Indigenous African languages, furthermore, rarely underwent elaborate corpus planning during the colonial era. The cost of corpus planning for multiple languages has seemed prohibitive, particularly given low rates of African literacy (Albaugh 2014: 55–56). Lacking legal reference works, technical manuals, and educational materials in indigenous languages, newly independent African states typically retained colonial languages for legal, scientific, and educational purposes.

Campaigns for linguistic indigenization, which could be seen as a form of language revival, have taken place in several postcolonial African states, but indigenous languages have so far gained little ground against European colonial languages. Nigerian linguist Ayo Bamgbose (2003: 422) concluded that "while post-independence governments appear to be making language policy, most of the time they are only perpetuating colonial language policy". Yet English, the former imperial language of Nigeria, can paradoxically claim to be Nigeria's national language in one sense that no indigenous language can match: it is equally (un)familiar to speakers from all the country's many language communities. Igbo, a Niger-Congolese language spoken in Southeastern Nigeria, may

be indigenous to Nigeria, but those promoting its use would be espousing Biafran separatism, not Nigerian nationalism (Chiluwa 2012: 227–230). English, of course, has shallow roots as a spoken language in Nigeria's rural communities, but Nigerian nationalism also has shallow roots in rural areas, where tribal, religious, or other sub-state loyalties dominate.

Postcolonial states in Asia have more frequently adopted indigenous languages alongside the former colonial languages, not least because many Asian languages boast extensive literary traditions. Nevertheless, many Asian states have also inherited multilingual populations from European colonial times, or indeed from precolonial times. The politics of selecting an indigenous language to replace or supplement former colonial languages has often proved contentious. Independent Pakistan, for example, adopted two official languages: English and Urdu. Urdu was chosen partly because of prestigious associations with the struggle against British rule, but partly as a neutral lingua franca. Most speakers of Urdu actually live in India: Pakistan's first census in 1951 found that Urdu was the first language of only 3.3 % of Pakistan's population. Bengalis in East Pakistan, mostly unfamiliar with Urdu, protested its favored status and demanded the use of Bengali. Both processes contributed to Bangladeshi separatism. Indeed, the 1952 deaths of Bengali students protesting against Urdu is commemorated as "language day" in independent Bangladesh (Ayres 2009: 45, 73). Yet when the Pakistani government, chastened by Bangladeshi independence, subsequently made concessions to Sindhi, Urdu speakers rioted (Siddiqi 2012: 92–93). Balancing the aspirations of different linguistic groups requires good will and a willingness to compromise. Sometimes both are in short supply.

In postcolonial states, linguistic minorities often prefer the former colonial language to the indigenous language of ethnic rivals. In 1965, for example, the Indian government abandoned efforts to replace English with Hindi because the perceived threat of Hindi dominance had provoked resistance and political separatism in non-Hindi-speaking provinces, and particularly in Dravidian-speaking Tamil Nadu (Dasgupta 1970: 225–242). If linguistic separatism no longer appears a major threat to the Republic of India, the country's federal structure deserves much of the credit: each of India's 29 states and 7 union territories can specify its own official languages, and 22 languages enjoy official status.

Even in the absence of political violence, however, promoting indigenous languages may conflict with other state priorities. In 1959, the government of newly independent Malaysia declared Bahasa Melayu the national language. The Malaysian government promoted Bahasa Melayu in the educational system, and devoted significant resources to reference books, technical dictionaries, and other forms of corpus planning. Even though nearly half of Malaysia's population speaks Tamil or Chinese, economic success has mostly shielded independent Malaysia from ethnic conflict and linguistic separatism. Nevertheless, the Malaysian government still chose to abandon its campaign for Bahasa Melayu in 2002. Since then, Malaysia has promoted English-language education, viewing improved mastery of English as a surer path to economic growth and technological progress (Gill 2005).

6. Prospects for the future

Speculations about future linguistic nationalism cannot be documented: studying the past is easier than foreseeing the future. Given how dramatically the industrial revolution transformed linguistic politics, however, we should expect future economic transformations to have equally far-reaching consequences. Increasing economic interdependence, easy transcontinental transport, telecommunications technology, and other factors collectively described as 'globalization' have changed attitudes toward multilingualism. National states increasingly tolerate multilingualism, and particularly encourage citizens to learn prestigious languages of global commerce. English has acquired an unprecedented status as a global language (Crystal 2012; Northrup 2013), but Spanish, Chinese, Japanese, French, and Russian are widely learned as second languages.

In the future, globalization may create new forms of linguistic politics. Just as increased wealth created a literate constituency for national movements during the 19th and 20th centuries, rising education levels may in the future create a social basis for political structures more cosmopolitan than the nation-state. The European Union, for example, promotes bi-, tri-, and multilingualism, and a successful career in its bureaucracy generally requires both oral and written facility in several languages (Huguet and Lasagabaster 2007: 234–37). Indeed, several advocates of multilingualism have invested great hopes in the promotion of a European identity, imagined as supplementing and enhancing particular nations, and thus, by extension, any one national language (Rindler Schjerve and Vetter 2012: 161–64). If a class of people mastering several literary languages eventually forms a sizable percentage of the world's population, it may form the social basis for new forms of globalized, post-national politics.

Nevertheless, immigration has become an important and divisive political issue in precisely those rich countries best situated to create a social class of multilinguals. European states generally accommodate existing linguistic minorities, but the new urban multilingualism often additionally involves new linguistic minorities without historic roots or any local rural hinterland. Communities speaking Turkish in Germany, Berber and Arabic in France, and Punjabi and Urdu in Britain have created new administrative challenges (Extra and Yağmur 2004). Liberal attempts at accommodation struggle against a nativist backlash hostile to both immigrants and globalization. How such tensions can be resolved remains to be seen.

7. References

Albaugh, Ericka
2014 *State-building and multilingual education in Africa*. Cambridge: Cambridge University Press.

Andraschke, Peter & Helmut Loos (eds.)
2002 *Ideen und Ideale: Johann Gottfried Herder in Ost und West*. Freiburg im Breisgau: Rombach.

Ayres, Alyssa
2009 *Speaking like a state: Language and nationalism in Pakistan*. Cambridge: Cambridge University Press.

Bamgbose, Ayo
2003 A recurring decimal: English in language policy and planning. *World Englishes* 22(4). 419–431.
Bauschatz, John
2013 *Law and enforcement in Ptolemaic Egypt*. Cambridge: Cambridge University Press.
Baxter, Alan
1992 Portuguese as a pluricentric language. In Michael Clyne (ed.), *Pluricentric languages: Differing norms in different nations*, 11–44. Berlin: Walter de Gruyter.
Bentahila, Abdelal & Eirlys Davies
1993 Language revival: Restoration or transformation? *Journal of Multilingual and Multicultural Development* 14(5). 355–374.
Boltz, William
1994 *The origin and early development of the Chinese writing system*. New Haven: American Oriental Society.
Brown, Cecil
1999 *Lexical acculturation in Native American languages.* Oxford: Oxford University Press.
Cardinal, Linda
2015 State tradition and language regime in Canada. In Linda Cardinal & Selma Sonntag (eds.), *State traditions and language regimes*, 29–43. Montreal: McGill-Queens.
Chiluwa, Innocent
2012 Social media networks and the discourse of resistance: A sociolinguistic CDA of Biafra online discourses. *Discourse and Society* 23(3). 217–244.
Çolak, Yilmaz
2004 Language policy and official ideology in early republican Turkey. *Middle Eastern Studies* 40(6). 67–91.
Crystal, David
2012 *English as a global language*. Cambridge: Cambridge University Press.
Dasgupta, Jyotirindra
1970 *Language conflict and national development: Group politics and national language policy in India*. Berkeley: University of California.
Deacon, Bernard
2006 Cornish or Klingon? The standardization of the Cornish language. *Cornish Studies* 14(1). 13–23.
DeFrancis, John
1984 *The Chinese language: Fact and fantasy.* Honolulu: University of Hawaii Press.
Despalatović, Elinor Murray
1975 *Ljudevit Gaj and the Illyrian movement*. Boulder: East European Monographs.
Doğançay-Aktuna, Seran
2004 Language planning in Turkey: Yesterday and Today. *International Journal of the Sociology of Language* 165. 5–32.
Extra, Guus & Kutlay Yağmur
2004 *Urban multilingualism in Europe: Immigrant minority languages at home and school.* Bristol: Multilingual Matters.
Fishman, Joshua A.
1997 *In praise of the beloved language: A comparative view of positive ethnolinguistic consciousness*. Berlin: De Gruyter.
Frawley, William, Kenneth Hill & Pamela Munro
2002 *Making dictionaries: Preserving indigenous languages of the Americas*. Los Angeles: University of California Press.
Gill, Saran Kaur
2005 Language policy in Malaysia: Reversing direction. *Language Policy* 4. 241–60.

Greenberg, Robert
2004 *Language and identity in the Balkans: Serbo-Croatian and its disintegration*. Oxford: Oxford University Press.
Haig, Geoffrey
2004 The invisibilisation of Kurdish: The other side of language planning in Turkey. In Stephan Conermann & Geoffrey Haig (eds.), *Die Kurden: Studien zu ihrer Sprache, Geschichte und Kultur*, 121–150. Hamburg: EB-Verlag.
Hannas, William
1997 *Asia's orthographic dilemma*. Honolulu: University of Hawaii.
Hassanpour, Amir
1992 *Nationalism and language in Kurdistan, 1918–1985*. San Francisco: Mellen Research University Press.
Haugen, Einar
1966 Dialect, Language, Nation. *American Anthropologist* 68(4). 922–935.
Haugen, Einar
1987 *Blessings of Babel: Bilingualism and language planning, problems and pleasures*. Berlin: De Gruyter.
Herder, Johann Gottfried von
1784–1791 *Ideen zur Philosophie der Geschichte der Menschheit*, four volumes. Riga: Johann Hartknoch.
Hobsbawm, Eric
1992 *Nations and nationalism since 1780: Programme, myth, reality.* Cambridge: Canto.
House, Deborah
2002 *Language shift among the Navajos: Identity, politics, and cultural continuity.* Tuscon: University of Arizona Press.
Hroch, Miroslav
1985 *Social preconditions of national revival in Europe: A comparative analysis of the social composition of patriotic groups among the smaller European nations.* Cambridge: Cambridge University Press.
Hroch, Miroslav
1994 The social interpretation of linguistic demands in European national movements. *EUI Working Paper, EUF* 94(1). Florence: European University Institute.
Huguet, Ángel & David Lasagabaster
2007 The linguistic issue in some European bilingual contexts: Some final considerations. In David Lasagabaster & Ángel Huguet (eds.), *Multilingualism in European bilingual contexts: Language use and attitudes*, 234–252. Bristol: Multilingual Matters.
Jensen, Eva Skafte
2016 The history of nominative -er in Danish and Swedish: A case of exaptation? In Muriel Norde & Freek Van de Velde (eds.), *Exaptation and language change*, 287–316. Amsterdam: John Benjamins.
Kamusella, Tomasz
2009 *The politics of language and nationalism in modern Central Europe*. London: Palgrave.
Kloss, Heinz
1967 *Abstand* languages and *Ausbau* languages. *Anthropological linguistics* 9. 90–101.
Laitin, David
1992 *Language repertoires and state construction in Africa*. New York: Cambridge University Press.
Lieber, Michael
1997 Murky minds, many muddles: The concept of race. In Larry Naylor (ed.), *Cultural diversity in the United States*, 53–74. Westport: Greenwood.

McRae, Kenneth
1983 *Conflict and compromise in multilingual societies: Switzerland*. Waterloo: Wilfrid Laurier University Press.
Messing, Jacqueline & Refugio Nava Nava
2016 Language acquisition, shift, and revitalization processes in Latin America and the Caribbean. In Serafín Coronel-Molina & Teresa McCarty (eds.), *Indigenous language revitalization in the Americas*, 76–96. New York: Routledge.
Michalowski, Piotr
2008 Sumerian. In Roger Woodard (ed.), *The ancient languages of Mesopotamia, Egypt and Aksum*, 6–45. Cambridge: Cambridge University Press.
Millar, Robert
2010 *Authority and identity: A sociolinguistic history of Europe before the modern age.* Basingstoke: Palgrave-Macmillan.
Miller, Nicholas
1995 Two strategies in Serbian politics in Croatia and Hungary before World War I. *Nationalities Papers* 23(2). 327–352.
Muljačić, Žarko
1993 Standardization in Romance. In Rebecca Posner & John N. Green (eds.), *Bilingualism and linguistic conflict in Romance*, 71–77. Berlin: De Gruyter.
Nakazawa, Takuya
2015 The making of 'Montenegrin language': Nationalism, language planning, and language ideology after the collapse of Yugoslavia (1992–2011). *Südosteuropäische Hefte* 4(1). 127–141.
Neustupný, Jiří
1970 Basic types of language treatment. *Linguistic Communications* 7. 78–117.
Norman, Jerry
1988 *Chinese*. Cambridge: Cambridge University Press.
Northrup, David
2013 *How English became the global language*. London: Palgrave.
Ohlsson, Stig Örhan
2005 Language contact in the 16th, 17th and 18th centuries: The Kingdom of Sweden. In Oskar Bandle (ed.), *The Nordic languages*, vol. 2, 1361–1368. Berlin: De Gruyter.
Okey, Robin
2004 Serbian, Croatian, Bosnian? Language and nationality in the lands of former Yugoslavia. *East European Quarterly* 38(4). 419–442.
Paulston, Christina Bratt, Pow Chee Chen & Mary Connerty
1993 Language regenesis: A conceptual overview of language revival, revitalisation and reversal. *Journal of Multilingual and Multicultural Development* 14(4). 275–286.
Peersman, Catharina, Gijsbert Rutten & Rik Vosters (eds.)
2015 *Past, present and future of a language border: Germanic-Romance encounters in the Low Countries*. Berlin: de Gruyter.
Peterson, Glen
2011 *The power of words: Literacy and revolution in South China, 1949–95.* Vancouver: University of British Columbia.
Peti-Stantić, Anita
2009 Restandardizacija (hrvatskoga) standardoga jezika. In Lada Badurina, Ivo Pranjković & Josip Silić (eds.), *Jezični varijeteti i nacionalni identiteti*, 71–82. Zagreb: Disput.
Pollock, Sheldon
2006 *The language of the gods in the world of men: Sanskrit, culture, and power in premodern India*. Los Angeles: University of California.

Radner, Karen (ed.)
2014 *State correspondence in the ancient world: From New Kingdom Egypt to the Roman Empire*. Oxford: Oxford University Press.
Rindler Schjerve, Rosita & Eva Vetter
2012 *European multilingualism: Current perspectives and challenges*. Bristol: Multilingual Matters.
Safran, William
2010 Political science and politics. In Joshua A. Fishman & Ofelia García (eds.), *Handbook of language and ethnic identity*, vol. 1, 49–69. Oxford: Oxford University Press.
Siddiqi, Farhan
2012 *The politics of ethnicity in Pakistan: The Baloch, Sindhi and Mohajir ethnic movements*. London: Routledge.
Sonntag, Selma
1991 *Competition and compromise amongst elites in Belgian language politics* (Plurilingua XII). Bonn: Dümmler.
Stanford, Nichole
2016 *Good God but you smart! Language prejudice and upwardly mobile Cajuns*. Boulder: University Press of Colorado.
Šubarić, Lav
2015 From the aftermath of 1784 to the Illyrian Turn: The slow demise of the official Latin in Croatia. In Gábor Almási & Lav Šubarić (eds.), *Latin at the crossroads of identity: The evolution of linguistic nationalism in the Kingdom of Hungary*, 191–217. Leiden: Brill.
Tomić, Olga Mišenka
1992 Macedonian as an *Ausbau* language. In Michael Clyne (ed.), *Pluricentric languages: Differing norms in different nations*, 437–454. Berlin: Mouton de Gruyter.
Troch, Pieter
2014 From 'and' to 'either/or': Nationhood in Montenegro during the Yugoslav twentieth century. *East European Politics and Societies* 28(1). 25–48.
van Goethem, Herman
2010 *Belgium and the monarchy: From national independence to national disintegration*. Brussels: UPA.
Versteegh, Kees
1984 *Pidginization and creolization: The case of Arabic*. Amsterdam: John Benjamins.
Versteegh, Kees
2014 *The Arabic language*. Edinburgh: Edinburgh University Press.
Weber, Eugen
1976 *Peasants into Frenchmen: The modernization of rural France*. Stanford: Stanford University Press.
Yuming, Li
2015 *Language planning in China*. Berlin: De Gruyter.

Alexander Maxwell, Wellington (New Zealand)

30. Language politics, policy, and planning

1. Introduction
2. Language politics, policy, and planning: conceptualizations and interrelations
3. Subcategories of LPP
4. Actors of LPP in the context of language domains
5. Aims, ideologies, and rights in LPP
6. The relevance of analyzing LPP in contact linguistics
7. Outlook
8. References

1. Introduction

Research on language politics, policy, and planning is of importance to contact linguistics, since political relations between groups of language users, the way in which the use of language(s) is organized, and how language issues are politicized fundamentally shape the political and social conditions under which language varieties are in contact. This chapter first provides a short sketch of how language policy, planning, and politics have so far been conceptualized. Major subfields will be discussed, and then relevant actors and factors in these processes will be introduced. At the end, these aspects will be discussed from a contact linguistic perspective and summarized in a graphic visualization.

2. Language politics, policy, and planning: conceptualizations and interrelations

Language politics, language policy and language planning are closely related. Language policy and planning are two sides of the same coin, with *policy* traditionally referring to more conceptual strategies and activities, and *planning* to direct intervention into language practices, i.e. how governments and other institutions actively try to influence the way in which languages are used. In more recent literature, the term *language policy* has become dominant as an overarching term for all kinds of activities and underlying motivations for policies dealing with language; alternatively, language policy and planning (LPP) has also become common, since it often proves difficult to neatly distinguish policy from planning activities (Hornberger 2006: 25).

A somewhat distinct tradition, even though closely related to LPP research, is language politics and the politics of language (see Wodak and Forchtner [2018] for an overview). These concepts denote how language is politicized, how language(s) and their speakers are used for establishing distinctions between groups, and how politics are organized around them, e.g. in discourses of nationalism, in which language is a fundamental marker of in- versus out-groups (Schmid 2001). The difference between these two is a matter of focus: whereas the politics of language reflects the instrumentalization of language for specific ideological purposes, language politics relates to the way in which political actors try to influence existing policies based on language. Yet LPP, language politics, and the politics of language are interlinked, since political questions

https://doi.org/10.1515/9783110435351-030

discussed along the lines of groups of speakers usually involve issues surrounding the status of different varieties, i.e. of acceptable versus unacceptable forms in specific situations. As such, the politics of language shares with LPP debates on ideologies, language practices by groups of speakers, and the ways in which language regimes are (consciously and subconsciously) organized. In addition, the politics of language and the organization of societies along language lines also involve different topics, e.g. when issues of power and of participation in society are decided according to peoples' language competences, or even according to their L1s. At the same time, language politics, policy, and planning are intertwined with other concepts: linguistic culture (Schiffman 1996: 8) denotes a "complex of values, beliefs, myths etc. concerned with language", which provides the setting for LPP, whereas language cultivation (Nekvapil 2008) focuses on aesthetic values, and language critique denotes evaluations of varieties and specific forms at the meta-level (Kilian, Niehr, and Schiewe 2010).

Research on specific cases of political influence on language(s) started to develop into a separate field of linguistics after World War II. It gained theoretical momentum in the 1960s and 1970s. Early understandings of LPP by e.g. Haugen (1959) focused on state-centered activities, mostly regarding variation within a language. Rubin and Jernudd (1971: xviii), in a famous definition, conceptualized language planning as "deliberate language change […] in the systems of language code or speaking or both that are planned by organizations that are established for such purposes or given a mandate to fulfil such purposes".

In the course of the academic debates which followed, the focus widened considerably, as exemplified by Cooper's (1989: 45) definition, according to which language planning denotes "deliberate efforts to influence the behavior of others with respect to the acquisition, structure, or functional allocation of their language codes". The debate now included broader aspects of linguistic behavior by far; it also made reference to the relationship between different language varieties. In the 1990s, LPP studies followed the tendency toward investigating language conflict and contact, in line with the 'critical turn' focusing on ideologies and power relations (Tollefson 2002).

More recent understandings of LPP therefore largely agree that perceptions beyond deliberate, active intervention into a linguistic situation need to be considered. Spolsky (2004, 2009) stipulates that language policy consists of three parts: language management (i.e. LPP in the traditional sense), language beliefs (i.e. conscious and subconscious views that influence which values and functions are assigned to specific varieties and variants), and language practices (i.e. what people actually do with language). This understanding of language policy suggests that essentially any type of linguistic behavior is the expression of a language policy: by preferring a specific variety or variant, every language user makes a statement, which in turn may influence the practices and beliefs of others. The ideologies expressed in this way reflect the status of varieties and their speakers in a political or societal context, as well as power relations. In any analysis, it is therefore important not only to consider *overt* policies, as expressed by language laws, institutional language plans, or the like, but also the *covert* side − language behavior which reflects policies and attitudes, as well as the 'hidden agendas' (Shohamy 2006) of language policy-making.

Somewhat aside from these theoretical developments, which looked at LPP mostly from a macro-perspective, Language Management Theory adopted a micro-approach (Jernudd and Nekvapil 2012; Jernudd and Neustupný 1987; Nekvapil 2012, 2015), sug-

gesting a circle-shaped model for analyzing active intervention into linguistic behavior: a specific linguistic situation is evaluated and discussed; if it is considered to be a problem, a decision may evolve which tries to change the linguistic behavior of the individuals or the community of linguistic practices in question; after a while, the situation is reevaluated.

Today, all of these approaches continue to be influential, but their labeling as 'classical language planning theory' (see Taylor-Leech [2016] for an overview of different phases) indicates that more contemporary approaches have developed. These derive from different fields, for instance, ethnographic and discursive approaches (Johnson 2013; McCarty 2011), i.e. observation- and interview-based analyses of interaction on the ground, as well as studies of language policy documents, all of which aim at analyzing underlying social orders. This LPP school is rooted in the framework of 'critical' social sciences, which aim at understanding ideologies, power structures, and social realities. Hult and Johnson (2015) collect the most important research approaches, including, for example, discourse analysis, media studies, language attitudes, and nexus analysis. Yet more traditional approaches also continue to be of importance, including the analysis of state-driven LPP (e.g. of official language regimes) as manifested in language laws.

It is in this combination of more traditional and more recent approaches where the connection between LPP and the politics of language also becomes apparent − a connection which, as Darquennes (2011: 557) remarks, has only fairly recently gained more systematic attention in academia. Kymlicka and Patten's (2001) comprehensive overview of relationships between political ideologies and their stances on language(s), and McRae's (1983, 1986, 1997) subsequent analyses of language-related political cleavages in Belgium, Switzerland, and Finland, can be seen as pioneering works.

3. Subcategories of LPP

Whereas recent writings frequently stress how multilayered and interwoven LPP activities (and their analysis) are, it is still beneficial to return to several traditional distinctions between types of language policies, since these continue to be used as important reference points (Hornberger 2006: 28−29; Kaplan and Baldauf 1997; Marten 2016: 27; Spolsky 2012). A common categorization going back to Kloss (1969) distinguishes between corpus and status planning. These two main strands of language policy-making are often separated in practice (e.g. public bodies and language agencies, such as the Académie Française, which develop the corpus of a language, in contrast to laws, which determine the official functions of different varieties), even though this distinction has been questioned (Fishman 2006). In brief, status planning considers the functions of different varieties in society (e.g. which languages are allowed to be used in official contexts), whereas corpus planning considers the choice of lexical, grammatical, and phonetic forms within a variety. A main issue is standardization, starting with Haugen's (1966) seminal work on the relationship between the two standardized Norwegian varieties (Bokmål and Nynorsk), and efforts to unite them by creating 'Samnorsk' as a common standard.

Other important categories are acquisition and usage planning (Cooper 1989; Lo Bianco 2005). Acquisition planning denotes processes that influence which varieties are acquired

as L1, L2, or foreign languages, most notably in formal education from kindergarten through university. Usage planning provides usage opportunities, e.g. by creating active offers which encourage speakers to expand the domains in which a variety is used, often in contexts of minority language policies.

Prestige planning (introduced by Haarmann [1990] as an addition to the status-corpus dichotomy) and discourse planning (Lo Bianco 2005) are terms which were added to the debate at later stages. Prestige planning affects views of specific varieties, e.g. when spreading understanding for the use of a minority language or − in nation-building contexts − for processes in favor of a common national language, e.g. when promoting Swahili in Tanzania after independence, replacing English, but to the detriment of less widely used languages (Batibo 1992). Prestige planning is related to status and acquisition planning, e.g. by demonstrating that a specific variety considered unsuitable in formal situations (e.g. in communication with public authorities) may indeed be used in such contexts, and by providing opportunities for acquiring the language either from scratch, or for expanding knowledge in more prestigious and specialized domains. Discourse planning denotes attempts at influencing language use at the meta-level, both with regard to LPP debates as such, and to discussions about values assigned to specific varieties and forms. This concept was introduced to LPP theory, not least because of "skepticism about the real impact of scholarly contributions to language planning [which aims] to explore closer conceptual links with policy analysis scholarship that theorizes power and with the various branches of discourse studies" (Lo Bianco 2005: 263). In this way, contemporary LPP research incorporates ideological stances and power relations which shape the political systems in which debates on LPP take place.

In all this, language policies may be categorized as symbolic, as opposed to practical. Symbolic steps include, for instance, the presence of a language on signs in the so-called 'linguistic landscape', i.e. the display of a language on visible signs in a public space (Shohamy and Gorter 2008), or the use of a variety in a speech by an activist or a politician. Such measures assign symbolic value to a variety, but they do not change the language behavior of other persons as such. This stands in contrast to active practical offers to e.g. acquire a language in education or use it in a company. In practice, it is often a combination of a symbolic recognition with practical steps which encourage the use of a specific variety, or which provide opportunities for influencing a linguistic environment.

4. Actors of LPP in the context of language domains

As has become apparent, many individuals and groups interact in determining which discourses on language prevail in a given society, and what implications these have for the status and use of varieties in contact in a linguistic ecosystem, as well as for the development of the standardized corpus of a variety. It is therefore useful to more closely examine groups of actors in specific contexts, in particular with regard to contemporary views of LPP which respect the perceptions of the various strata of society.

Early studies mostly investigated the state and semi-state institutions − governments, ministries, language academies, or councils − which openly shape language policies. Closely related to state policies are usually educational policies and state-funded kinder-

gartens, schools, and universities, which follow official rules in their choices about media of instruction and languages offered for acquisition. Private institutions are at least subject to specific laws.

Yet it is commonly accepted today that nongovernmental institutions, private companies, and individuals also conduct individual language policies − partly through overt statements, and partly through *de facto* practices based on underlying beliefs (the tendency to broaden the understanding of language policy has become apparent in titles such as *Language Policy Beyond the State* by Siiner, Koreinik, and Brown [2017]). A common distinction is made between 'top-down' and 'bottom-up' (with possible intermediate levels) policies and practices, which may or may not be juxtaposed, even though Johnson (2013) remarks that successful LPP usually requires that players from different levels have similar aims and interact with each other. Most notable bottom-up actors are activists, who support one variety (e.g. in minority language contexts), but nonprofit organizations as well as private companies essentially conduct (direct or indirect) language policies with regard to language choice in communication, or skills required by employees (note that this also applies to standard varieties as opposed to dialects or sociolects). Likewise, individuals at the private level also have conscious and subconscious policies, reflected in the development of research on family language policies in recent years (King, Fogle, and Logan-Terry 2008; Schwartz and Verschik 2013). Families make decisions about which varieties to teach their children (most prominently discussed in the context of the one-parent-one-language model in bilingual families), but also which forms are considered 'good' or 'bad' language when raising children, and when choosing languages in education.

Another important tool of categorization − used in LPP research and by many language planners in practice − are domains (Spolsky 2004). Even if no individual acts in only one linguistic domain (and even though holistic approaches attempt to overcome separations), LPP is frequently organized by the domain approach, which has been referred to as one of the major approaches to LPP (e.g. languages in education, public administration, the health system, courts and police, cultural and heritage organizations, and the media [Baldauf 2012]). Among the most important official domains are lawmaking (see Turi [2012] on a classification of language legislation), public bodies, and education (Tollefson 2013). More private domains include the economy and businesses, but also culture or heritage activities, as well as religious institutions. Of particular importance are also media (both traditional types and more recent innovations), as they help to spread certain varieties while also shaping discourses on language(s).

A field which stands somewhat outside the traditional purview of LPP is the field of international language policies. National language institutes (Ammon 2014) support their respective national languages both within and outside of their nation states. Exterior language policies enacted by embassies and by cultural and language institutes, such as the British Council or Goethe-Institut, aim at the promotion of national languages outside their core areas, including activities at all levels of education (e.g. by financing lecturers at universities abroad). Language, cultural, educational, and economic aims often go hand in hand in such processes. International language policies may also take shape as policies in international institutions or informal networks. Organizations such as the UN or the EU have long LPP traditions (McEntee-Atalianis 2015), both with regard to internal practices (as official or working languages, or languages of communication and publications), and to support for specific languages, e.g. through minority language pro-

tection or technological support. Informal international networks exist not least in academic circles on linguistics; often, they take an interdisciplinary shape.

5. Aims, ideologies, and rights in LPP

In line with the distinction between overt and covert policies, the aims of LPP may be explicitly formulated in policy documents, or be implicitly understood. Nahir (1984) lists categories of LPP aims, e.g. standardization (typical of corpus planning), maintenance and revitalization (of endangered languages), terminological unification, and lexical modernization. Vikør (1994) classifies such principles, namely: language-internal principles (e.g. simplification of phonological and/or morphological structures), the relationship to other varieties (e.g. purism), and the relationship between a variety and its speakers (for instance, the so-called 'majority principle', which accepts the most common variants as the norm, and the 'revolt principle', which supports forms which may be considered provocative by many users, e.g. feminist language policies which aim at spreading awareness of gender inequalities).

At the core of the politics of language is the relationship between language and general political ideologies. Related to Spolsky's (2004) 'language beliefs', ideologies are stances shared by larger sections of society. Cobarrubias (1983) noted four standard ideologies: assimilation, pluralism, vernacularization, and internationalization. Ruíz (1984) uses the concept of 'language orientations' to categorize whether a language is considered a problem, a right, or a resource. Vikør (1994) applies to language the concept of liberalism, i.e. a lack of interference which leaves language use to market forces (cf. also the critique by Robichaud and De Schutter [2012]). Other underlying ideologies include traditionalism (i.e. traditional forms are considered valuable *per se*), and nationalism.

The focus on ideologies gained momentum during the 'critical turn' (Blommaert 1999), and is at the core of many contemporary LPP studies, which explicitly aim to identify ideologies and their consequences for users of specific varieties (Tollefson 1991). In particular, with regard to nationalism, language policies of many nation states favor a national language, to the detriment of both traditional minorities and migrant communities (see Gogolin [2008] on Germany; Schiffman [1996] on France). In the educational field, but also in other domains, official ideologies are reinforced, e.g. with regard to the question of which varieties are accepted in communication with a public body. Factors which were investigated in early LPP research, such as nationalism (Fishman 1972), continue to be fundamental for many LPP frameworks; yet nationalism has been sidelined by newer major ideologies, such as globalization (Wright 2016). In the sense of a 'hidden agenda', language questions are frequently used to deny speakers of the 'wrong' language, or of nonstandard variants, access to influential positions in society (for an overview of ideological debates around English versus other languages in the US throughout the country's history, for instance, see Ramsey 2012).

A case in point is the Baltic states, where top-down LPP influences the status of the national and minority languages, but where after the end of the Soviet Union language shift reversal from Russian towards Estonian, Latvian, and Lithuanian, respectively, stands in contrast to defenders of Russian in a human rights' context: language is being

politicized by both the Baltic governments as part of the one-nation-one-language ideology, as well as by the politicians of Russian (Ozolins 2019). A classic example of a monolingual ideology is France, where the state actively engages in corpus planning and also provides an extensive framework for political decisions, most notably through language acts, for enshrining the dominance of French, e.g. in the media or education *vis-à-vis* both minority (such as Breton, Corsican, or Alsatian) and international (i.e. mostly English) languages. In multilingual countries, societies and territories are frequently split by language, including regular renegotiation of the principles of linguistic coexistence (e.g. Belgium, Switzerland, and Finland). Yet subunits of states also use strong ideological claims in order to demand an increase in status, acquisition, usage, and/or prestige of minority languages which are constitutive for these regions, and which may be more (e.g. Quebec and Catalonia) or less (e.g. Wales and Friesland) connected to political separatism.

Nationalism and other factors reinforced by a state are often supported by the beliefs and practices of its majority population, but they also frequently stand in sharp contrast to the practices of other speech communities. As a counterweight, individuals and institutions at the micro-level seek out individual ways to facilitate more multilingual practices. In LPP research, investigating such opposing processes, and their underlying ideologies, is fundamental for understanding how language contact creates social conflicts stemming from e.g. socioeconomic status (sociolects, urban dialects), center versus periphery (i.e. traditional rural dialects), or distinctions with regard to the languages of migrants and/or of autochthonous minorities. A specific field of LPP research is therefore minority language policy (Fishman 2001), which often aims at improving the chances of language revitalization, or a reversal of language shift (famous examples being, for instance, Welsh, Maori, and Native American languages such as Navajo; see Hinton and Hale 2001). Varieties may be strengthened through improved status (e.g. by making them official, or co-official, in a state and/or region); yet status by itself does not guarantee language survival. Therefore, acquisition and usage planning are often at the core of activities, supplemented by attempts to raise prestige (among both the minority and the majority populations) and to influence discourses on the desirability of the survival of a specific variety, often in relation to the specific identity of a region.

Debates on minority languages are frequently connected to general issues of language rights. These are subject to international treaties and other regulations, such as the UN Charter, which provides a very broad antidiscrimination regime on the basis of languages. More specific is the 1992 UN Declaration on the Rights of Persons Belonging to National or Ethnic, Religious and Linguistic Minorities (Orlin 2015). In the European context, documents by the Council of Europe protect languages and their speakers, including in particular the 1992 Charter of Regional or Minority Languages, and the 1995 Framework Convention for the Protection of National Minorities (Grin 2003).

6. The relevance of analyzing LPP in contact linguistics

LPP relates to language contact in numerous ways (see Darquennes [2013] for an overview). Corpus planning is the field which reacts most visibly to language contact. Through norm-setting processes, such as the acceptance (or refusal) of loan words or of

grammatical interference in a variety, the degree to which language contact is 'allowed' to influence a variety is regulated. This relates most notably to official standardization processes, e.g. the sanctioning of dictionaries or teaching materials. Darquennes (2015) stresses the *Ausgleich* function of LPP measures − active intervention which seeks a compromise between different views − as one possible way out of a (potential) language conflict. Successful intervention depends, however, on both institutional support and a positive social climate; implementation of such measures has therefore been called the "Achilles heel" (Deumert and Vandenbussche 2003: 7) of standardization. Edwards (2012) stresses that language agencies such as the Accademia della Crusca in Italy may indeed contribute to such an *Ausgleich*, while also at the same time enshrining tensions between elitist norms and real-life use. Latvia and Norway (Bull 1993) are examples of states which have reacted to contact-induced language change with active corpus planning, e.g. by suggesting neologisms as replacements for lexical borrowings. Yet speech communities also react to language contact 'from below' by incorporating loan words into a variety, through practices such as code-switching or translanguaging in multilingual settings. Here also, practices reflect language beliefs about which kind of influence is considered appropriate in which functions.

The coexistence of different varieties in contact also leads to questions about their status. LPP reacts to such processes by reinforcing the dominance of one variety (which usually leads to an increase in tensions), or by creating spaces for different varieties, based on the idea of equality of languages and respect for the wishes of different groups of speakers. Status questions include, for instance, the labeling of 'desired' and 'undesired' languages in education, but also defining whether a regional variety is a language in its own right, or rather a sub-variety of a standard language (e.g. Latgalian in Latvia, or Scots in Scotland). Prestige planning and usage planning are often related to status planning, e.g. in discourses about which varieties have more prestige and should therefore be given more usage opportunities.

In acquisition planning, language contact may also result in conflict-laden situations, in which LPP processes (and underlying ideologies) play an important role. In education, elements of the 'wrong' variety are often systematically abolished, in particular in traditional nation-states, such as Germany, where the 'monolingual habitus' requires standard proficiency for educational success as well as for access to prestigious domains (Gogolin 2008). In societies with a tradition of accepted multilingualism, language contact may be considered less problematic. Yet discourses on standards also frequently determine which linguistic repertoires are required for societal success, and politicize contact phenomena. Loan words, or practices such as translanguaging, for instance, are considered by some ideological stances to 'threaten' a language's 'purity' − which may be incorporated into educational standards.

Taking into consideration the multitude of aspects in LPP research discussed in this chapter, the following graphic visualization *Holistic Ecolinguistic Model for the Analysis of Language Policy* (*HEMALP*) suggests a way to summarize the most relevant structures and mechanisms within a geographical or political unit (Marten 2016), including conflict between different varieties and their users. It draws on previous graphic models, such as Kaplan and Baldauf's (1997) ecolinguistic tool for analyzing language regimes. Kaplan and Baldauf's (1997) idea was that different varieties in contact with each other shape the ecolinguistic reality in which varieties and speakers develop. These are influenced by speakers' practices, by overt as well as covert values and ideologies, and last but not

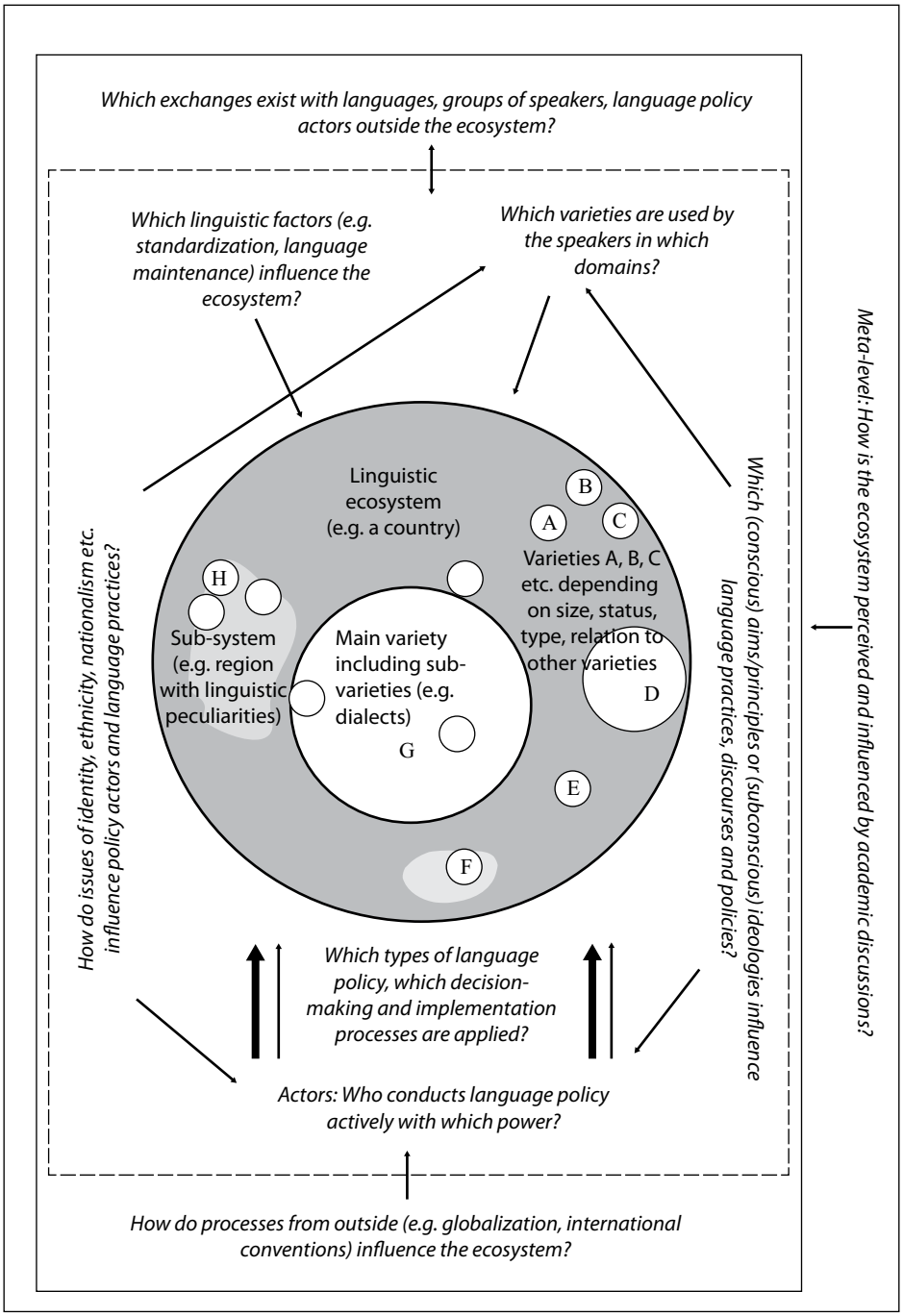


Fig. 30.1: Holistic Ecolinguistic Model for the Analysis of Language Policy (HEMALP), adapted from Marten (2016)

least, by status, corpus, usage, acquisition, and prestige planning activities. At the same time, the model incorporates contact-induced linguistic processes, such as structural changes and language revitalization.

With the benefit of more than two decades of research and insight after Kaplan and Baldauf (1997), HEMALP incorporates more recent academic debates. The model should be read in the following way: the main circle in the middle represents a linguistic ecosystem, i.e. a geographical or political unit (a state, region, city, or other); the small circles represent different varieties and their speakers, and the ways in which they are in contact (and possibly in conflict) with each other. Various actors, through their active LPP measures, influence the ecosystem: ideologies shape practices, which result in contact phenomena and status questions; rules and discourses from outside the ecosystem influence these active interventions, practices, and beliefs; aspects of language politics are part of the discourses on languages, but also of ideologies. Each variety is potentially influenced by LPP activities for purposes of identity formation, or of separating groups of society according to language, which results in disparate functions for varieties, which in turn influences practices, beliefs, discourses, and/or active LPP measures at the same time.

From the perspective of language contact, HEMALP displays which varieties are in contact with each other, and in which relationship (i.e. the size of the speech communities, as well as geographical or structural similarities). The model highlights factors which influence a contact situation, and shows how different LPP actors and speakers react to them − by direct intervention or as a speech community through their practices − and on which beliefs these reactions are based. HEMALP thereby allows for a contrastive perspective to compare the complexity of LPP factors in ecolinguistic systems in different political entities (e.g. states), or for a diachronic perspective to aid in understanding changes in a single region.

7. Outlook

In sum, this chapter has shown that language policy, planning, and politics are closely related, and are in their totality and complexity an important aspect of contact linguistics. Further research will be continuously necessary where new contact situations lead to new practices, fed by new (or reinvented) ideologies and other beliefs which − consciously or subconsciously − LPP actors will relate to, and turn into new interventions. Particularly fraught political (and linguistic) 'hot potatoes' at the time of writing, such as the crisis in Ukraine − in which the status of Russian plays an important role, both in the politics of language, where Russia equates language and ethnicity, and thereby justifies measures to 'protect' Russophones, and in LPP, where the question of the official status of Russian was one of the initial bones of contention in the crisis − show that both the birth and intensification of a language conflict, and its eventual resolution, may depend heavily on language politics, policy, and planning.

8. References

Ammon, Ulrich (ed.)
2014 *European national language institutions' attitudes and policies towards English as the European lingua franca* (Sociolinguistica 28). Berlin: de Gruyter.

Baldauf, Richard B.
2012 Introduction – language planning: Where have we been? Where might we be going? *Revista Brasileira de Linguística Aplicada* 12(2). 233–248.

Batibo, Herman M.
1992 The fate of ethnic languages in Tanzania. In Matthias Brenzinger (ed.), *Language death: Factual and theoretical explorations, with special reference to East Africa*, 85–98. Berlin: Mouton de Gruyter.

Blommaert, Jan (ed.)
1999 *Language ideological debates* (Language, Power and Social Process 2). Berlin & New York: Mouton de Gruyter.

Bull, Tove
1993 Conflicting ideologies in contemporary Norwegian language planning. In Ernst H. Jahr (ed.), *Language conflict and language planning*, 21–38. Berlin & Boston: de Gruyter Mouton.

Cobarrubias, Juan
1983 Ethical issues in status planning. In Juan Cobarrubias & Joshua A. Fishman (eds.), *Progress in language planning: International* perspectives, 41–85. The Hague: Mouton.

Cooper, Robert L.
1989 *Language planning and social change*. New York: Cambridge University Press.

Darquennes, Jeroen
2011 Minorities, language politics and language planning in Europe. In Bernd Kortmann & Johan van der Auwera (eds.), *The languages and linguistics of Europe: A comprehensive guide*, 547–560. Berlin & New York: de Gruyter.

Darquennes, Jeroen
2013 Current Issues in LPP research and their impact on society. *AILA Review* 26. 11–23.

Darquennes, Jeroen
2015 Language conflict research: A state of the art. *International Journal of the Sociology of Language* 235. 7–32.

Deumert, Ana & Wim Vandenbussche
2003 Standard languages: Taxonomies and histories. In Ana Deumert & Wim Vandenbussche (eds), *Germanic standardizations: Past to present*, 1–14. Amsterdam: Benjamins.

Edwards, John
2012 Language management agencies. In Bernard Spolsky (ed.), *The Cambridge handbook of language policy*, 418–436. Cambridge: Cambridge University Press.

Fishman, Joshua A.
1972 *Language and nationalism: Two integrative essays*. Rowley: Newbury.

Fishman, Joshua A. (ed.)
2001 *Can threatened languages be saved? Reversing language shift, revisited: A 21st century perspective.* Clevedon: Multilingual Matters.

Fishman, Joshua A.
2006 *DO NOT leave your language alone: The hidden status agendas within corpus planning in language policy.* New York: Lawrence Erlbaum Associates.

Gogolin, Ingrid
2008 *Der monolinguale Habitus der multilingualen Schule*. Münster et al.: Waxmann.

Grin, Francois
2003 *Language policy evaluation and Europe: The European Charter on Regional or Minority Languages*. New York: Palgrave Macmillan.

Haarmann, Harald
1990 Language planning in the light of a general theory of language: A methodological framework. *International Journal of the Sociology of Language* 86. 103–126.
Haugen, Einar
1959 *Planning for a standard language in modern Norway.* Anthropological Linguistics 1(3). 8–21.
Haugen, Einar
1966 *Language conflict and language planning: The case of modern Norwegian*. Cambridge, MA: Harvard University Press.
Hinton, Leanne & Kenneth Hale (eds.)
2001 *The green book of language revitalization in practice*. New York: Academic Press.
Hornberger, Nancy H.
2006 Frameworks and models in language policy and planning. In Thomas Ricento (ed.), *An introduction to language policy: Theory and method*, 24–41. Oxford: Blackwell.
Hult, Francis M. & David Cassels Johnson (eds.)
(2015) *Research methods in language policy and planning: A practical guide.* Oxford: Wiley-Blackwell.
Jernudd, Björn H. & Jiří Nekvapil
2012 History of the field: A sketch. In Bernard Spolsky (ed.), *The Cambridge handbook of language policy*, 16–36. Cambridge: Cambridge University Press.
Jernudd, Björn & Jiří V. Neustupný
1987 Language planning: For whom? In Lorne Laforge (ed.), *Proceedings of the international colloquium on language planning*, 69–84. Québec: Les Presses de l'Université Laval.
Johnson, David Cassels
2013 *Language policy.* Houndmills & New York: Palgrave Macmillan.
Kaplan, Robert B. & Richard B. Baldauf, Jr
1997 *Language planning: From practice to theory.* Clevedon: Multilingual Matters.
Kilian, Jörg, Thomas Niehr & Jürgen Schiewe
2010 *Sprachkritik: Ansätze und Methoden der kritischen Sprachbetrachtung*. Berlin: De Gruyter.
King, Kendall A., Lyn Fogle & Aubrey Logan-Terry
2008 Family language policy. *Language and Linguistics Compass* 2. 907–922.
Kloss, Heinz
1969 *Research possibilities on group bilingualism: A report*. Quebec: International Center for research on Bilingualism.
Kymlicka, Will & Alan Patten (eds.)
2001 *Language rights and political theory.* Oxford: Oxford University Press.
Lo Bianco, Joseph
2005 Including discourse in language planning theory. In Paul Bruthiaux, Dwight Atkinson, William Eggington, William Grabe & Vaidehi Ramanathan (eds.), *Directions in applied linguistics*, 255–264. Clevedon: Multilingual Matters.
Marten, Heiko F.
2016 *Sprach(en)politik: Eine Einführung*. Tübingen: Narr.
McCarty, Teresa L.
2011 *Ethnography and language policy.* New York: Routledge.
McEntee-Atalianis, Lisa J.
2015 Language policy and planning in international organisations. In Ulrike Jessner-Schmid & Claire J. Kramsch (eds.), *The multilingual challenge: Cross-disciplinary perspectives*, 295–322. Berlin: De Gruyter Mouton.
McRae, Kenneth D.
1983 *Conflict and compromise in multilingual societies: Switzerland.* Waterloo, Ontario: Wilfrid Laurier University Press.

McRae, Kenneth D.
1986 *Conflict and compromise in multilingual societies: Belgium.* Waterloo, Ontario: Wilfrid Laurier University Press.
McRae, Kenneth D.
1997 *Conflict and compromise in multilingual societies: Finland.* Waterloo, Ontario: Wilfried Laurier University Press.
Nahir, Moshe
1984 Language planning goals: A classification. *Language Problems and Language Planning* 8(3). 294–327.
Nekvapil, Jiří
2008 Language cultivation in developed contexts. In Bernard Spolsky & Francis Hult (eds.), *The handbook of educational linguistics*, 251–265. Oxford: Blackwell.
Nekvapil, Jiří
2012 From language planning to language management: J. V. Neustupný's heritage. *Media Komyunikeshon kenkyu / Media and Communication Studies* 63. 5–21.
Nekvapil, Jiří
2015 Language management theory as one approach in language policy and planning. *Current Issues in Language Planning* 17(1). 11–22.
Orlin, Theodore S.
2015 The death of languages; the death of minority cultures; the death of a people's dignity: Its implications for democracy and the commitment to human rights. In Heiko F. Marten, Michael Rießler, Janne Saarikivi & Reetta Toivanen (eds.), *Cultural and linguistic minorities in the Russian Federation and the European Union*, 47–79. Cham: Springer.
Ozolins, Uldis
2019 Language policy, external political pressure and internal linguistic change: The particularity of the Baltic case. In Sanita Lazdiņa & Heiko F. Marten (eds.), *Multilingualism in the Baltic states. Societal discourses and contact phenomena*, 29–55. London: Palgrave Macmillan.
Ramsey, Paul J. (ed.)
2012 *The bilingual school in the United States: A documentary history.* Charlotte: Information Age Publishing.
Robichaud, David & Helder De Schutter
2012 Language is just a tool! On the instrumentalist approach to language. In Bernard Spolsky (ed.), *The Cambridge handbook of language policy*, 124–145. Cambridge: Cambridge University Press.
Rubin, Joan & Björn H. Jernudd (eds.)
1971 *Can language be planned? Sociolinguistic theory and practice for developing nations*. Honolulu: East-West Center & University of Hawaii Press.
Ruíz, Richard
1984 Orientations in language planning. *NABE: The Journal for the National Association for Bilingual Education* 8(2). 15–34.
Schiffman, Harold F.
1996 *Linguistic culture and language policy.* New York: Routledge.
Schmid, Carol L.
2001 *The politics of language: Conflict, identity, and cultural pluralism in comparative perspective*. New York: Oxford University Press.
Schwartz, Mila & Anna Verschik (ed.)
2013 *Successful family language policy.* Dordrecht: Springer.
Shohamy, Elana
2006 *Language policy: Hidden agendas and new approaches*. London: Routledge.
Shohamy, Elana & Durk Gorter (eds.)
2008 *Linguistic landscape: Expanding the scenery.* London: Routledge.

Siiner, Maarja, Kadri Koreinik & Kara Brown (eds.)
2017 *Language policy beyond the state*. Cham: Springer.
Spolsky, Bernard
2004 *Language policy.* Cambridge: Cambridge University Press.
Spolsky, Bernard
2009 *Language management.* Cambridge: Cambridge University Press.
Spolsky, Bernard (ed.)
2012 *The Cambridge handbook of language policy.* Cambridge: Cambridge University Press.
Taylor-Leech, Kerry
2016 From practice to theory: The contributions of Richard B. Baldauf Jr. to the field of language policy and planning. *Current Issues in Language Planning* 17(1). 1–10.
Tollefson, James W.
1991 *Planning language, planning inequality.* New York: Longman.
Tollefson, James W.
2002 Limitations of language policy and planning. In Robert B. Kaplan (ed.), *The Oxford handbook of applied linguistics*, 416–425. Oxford: Oxford University Press.
Tollefson, James W. (ed.)
2013 *Language policies in education: Critical issues*, 2nd edn. New York: Routledge/Taylor & Francis.
Turi, Joseph G.
2012 Language law and language rights. *International Journal of Law, Language and Discourse* 2(4). 1–18.
Vikør, Lars S.
1994 *Språkplanlegging: Prinsipp og praksis*. Oslo: Novus.
Wodak, Ruth E. & Bernhard Forchtner (eds.)
2018 *The Routledge handbook of language and politics*. Abingdon: Routledge.
Wright, Sue
2016 *Language policy and language planning: From nationalism to globalisation*, 2nd edn. Basingstoke: Palgrave Macmillan.

Heiko F. Marten, Riga and Rēzekne (Latvia)

31. Majorities and minorities in language policy and language rights

1. Introduction
2. Defining (and contextualizing) linguistic minority status
3. Linguistic minorities: a sociological perspective
4. Linguistic minorities in international law
5. Conclusion
6. References

1. Introduction

The question of what language rights, if any, might be attributable to linguistic minorities in modern nation-states is a particularly vexed one. It has been with us since at least the

https://doi.org/10.1515/9783110435351-031

advent of modern nation-state formation, which most commentators agree was precipitated by the French Revolution of 1789 and the related rise of political nationalism (Gellner 1983; Hobsbawm 1990). A central feature of political nationalism is the principle of linguistic homogeneity − the confluence of citizenship and national belonging that requires the establishment of, and allegiance to, a common national language for all of a nation's citizens (Anderson 2006; May 2016; see also Chapter 29). Of course, a nation-state may have more than one national language. However, because of the principle of linguistic homogeneity, the majority of states still tend to have, at most, only a limited number of national languages. Even when multiple national languages are recognized (such as in the current South African constitution), in practice, only a few languages continue to predominate in the public domain (in the South African context, this is increasingly English; see Makalela 2017). A focus on linguistic homogeneity has resulted in a strong historical emphasis on acquiring nationally sanctioned languages, usually at the specific expense of other language varieties. This has in turn led to a related reluctance on the part of many nation-states to grant minority language rights, most commonly on the basis that it could potentially undermine a commitment to common citizenship and to related political stability (May 2012).

This lack of accommodation of minority language rights has also been a prominent feature of much international law. In the post-Second World War era, in particular, there has been a general renunciation of group-based rights in favor of universal individual human rights, most clearly articulated in the United Nations' 1948 *Declaration of Human Rights* (UNDHR). This has led to a related retrenchment of minority language rights, which are often framed on the basis of group entitlement. Over time, there has been some softening of this oppositional position. Key legislative developments, such as the *International Covenant on Civil and Political Rights* (ICCPR, 1966), the *European Charter for Regional or Minority Languages* (1992), the *Framework Convention for the Protection of National Minorities* (1994), and the *United Nations Declaration on the Rights of Indigenous Peoples* (2007), all include specific reference to minority language rights. However, the prospect of expanding minority language rights remains a challenging one in the current era.

In this chapter, I will explore the issues attendant to this question of language rights for linguistic minorities by first examining the notion of linguistic minority status itself, both sociologically and in relation to international (and national) law. In so doing, I emphasize that no determinative understanding of what constitutes a 'linguistic minority' can ever be reached, which must instead always be understood relationally, *vis-à-vis* majority language speakers, and contextually − socially, politically, and geographically. That being said, I posit that minority language rights nonetheless can (and should) be extended to linguistic minorities at the regional, national, and supranational levels. This better reflects the complex, dynamic, and fluid bi/multilingual repertoires of individual citizens. Contrary to still widely held notions of linguistic homogeneity, the recognition of minority language rights potentially bolsters, rather than undermines, wider political stability (de Varennes 1996; de Varennes and Kuzborska 2017, 2019; May 2018). It is also consonant with, I argue, the more accommodative, albeit still nascent, recognition of minority language rights in international law.

2. Defining (and contextualizing) linguistic minority status

Defining 'linguistic minority' status is a fraught process, and can only be sensibly pursued in relation to a deeply contextualized understanding of what any given minority/majority language relationship might comprise, including, crucially, any changes over time. That said, some broad parameters can be tentatively drawn. It is commonly accepted, for example, that the distinction between 'minority' and 'majority' languages is not primarily based on numbers (whereby fewer would equate to minority, and greater to majority), although numbers are inevitably a consideration. Rather, the minority/majority language distinction can best be described as a function of differences in status, power, influence, and entitlement among languages and their speakers. Majority language speakers have more of each; minority language speakers have less.

Majority languages are highly valued. They are reinforced by longstanding territorial and/or state language policies, allowing for their use in formal/public language domains, such as law, education, and public administration. Accordingly, majority languages are associated with communicative reach and wider social mobility. Minority languages are constructed in the opposite terms. They are most often associated with attenuated use and usefulness. This is because territorial or state policies have tended to limit them to low-status private domains (e.g. family, community), and they are thus also regarded as having little ongoing value. As a result, many minority language speakers eventually 'shift' to speaking the majority language in order to 'get ahead', dispensing with the minority language in so doing. As noted earlier, the minority/majority distinction has historically been the byproduct of the nationalism of the last few centuries, and the linguistically homogenous nation-state model to which it gives rise (Bauman and Briggs 2003; May 2012). However, it is also increasingly a feature at the supranational level, with the inexorable rise of English as the current *lingua mundi*, or global language, impacting other languages, including not only so-called minority languages, but national ones as well (Ives 2010; May 2014a).

It is this process of language shift and loss of minority languages that has been increasingly put into question by language rights' advocates. Language shift and loss have always occurred, but never as quickly, and never to this extent, as in this current, late-modern globalized era (see also Chapter 37). In so doing, questions have been raised as to why minority language speakers should not have at least some of the protections that majority language speakers enjoy (and take for granted). Yet this also inevitably raises the question as to what constitutes a 'linguistic minority group' − or, indeed, whether any group-based linguistic rights can be so recognized? Accordingly, in what follows, I first chart a broadly sociological understanding of who might comprise linguistic minorities. I then turn to related developments in international law concerning how to define linguistic minorities over time, along with the legal protections accorded to them − or not, as the case may be.

3. Linguistic minorities: A sociological perspective

From a sociological perspective, 'linguistic minorities' can be defined as language groups which are a numerical minority in a given state *and* which are also politically

non-dominant (Minority Rights Group 1997; see also http://minorityrights.org/directory/). As already indicated, this is not to suggest that minority/majority relationships are fixed, since they are, by definition, relative and relational − differing from one context to the next, as well as over time (Eriksen 2010). Nonetheless, linguistic minorities are usually characterized by a history of linguistic − as well as broader social, economic, and political − marginalization and/or exploitation within nation-states (Carens 2000). In order to make some distinctions among these minorities, in what follows, I draw on the social anthropologist Thomas Eriksen's (2010) useful attempts to develop a typology of such minority groups, albeit from a broader perspective than a solely linguistic one. Given the inevitable limitations of typologies, where appropriate, I also draw on a number of complementary categorizations. This attempt at a more nuanced categorization notwithstanding, the following distinctions should be seen primarily as useful heuristic devices, rather than as definitive and/or exhaustive categories. It should also be borne in mind that these minority groups are no more homogeneous than any other grouping, and accordingly, reflect significant *intra-* as well as *inter*group differences, along with a considerable degree of overlap.

3.1. Indigenous peoples

'Indigenous peoples' refer to aboriginal groups who are politically non-dominant, and who are not, or are only partially, integrated into the nation-state. They include such groups as Māori, Sámi, Australian Aboriginal peoples and Torres Strait Islanders, Native Americans, Hawaiians, and Inuit. These groups are associated historically with a non-industrial mode of production, and a stateless political system (Minority Rights Group 1997). The extreme disadvantages currently faced by many indigenous groups in modern nation-states are the result of colonization, and subsequent marginalization, within their own historic territories. Such historical processes have usually seen the expropriation of land, and the destruction (or near destruction) of their language(s) and traditional social, economic, and political practices. Accordingly, the predominant concerns of indigenous peoples are for separate (or at least more autonomous) political and cultural recognition *within* the nation-state, and where possible, for political and economic redress for past injustices.

These arguments have been most clearly articulated by indigenous peoples in various international law forums. Perhaps most notable here is the development of the *United Nations Declaration of the Rights of Indigenous Peoples* (2007), which I discuss in more detail in the next section. Such international developments have bolstered, and at times catalyzed, changes at the national level in favor of indigenous peoples, often with a specific focus on the extension of their language and education rights and provisions. Two brief examples suffice at this point. In Brazil, the adoption in 1988 of a new *Constituição* ('constitution') recognized for the first time *povos indígenas no Brasil* ('the indigenous peoples of Brazil'), of whom there are over 400,000. Article 231 specifically endorses indigenous social organization, customs, languages, beliefs, and traditions, along with the right of native title to their lands (Brasil 1996). That said, subsequent progress towards actualizing indigenous rights in Brazil has remained highly variable, with ongoing strong resistance from, in particular, entrenched mining and landholding interests (Poirier 2017).

Norway provides a similar, albeit more successful, example. After a century of enforcing a stringent 'Norwegianization' (read: assimilationist) policy towards the indigenous Sámi, their languages, and their culture, Norway moved in 1988 to revise its constitution in order to grant greater autonomy to Sámi. As the amendment to the Constitution stated: "It is incumbent on the governmental authorities to take the necessary steps to enable the Sámi population to safeguard and develop their language, their culture and their social life". The effects of this new amendment are most apparent in the regional area of Finnmark, in the northernmost part of Norway, where the largest percentage of Sámi peoples live. The formal recognition accorded to Sámi led to the subsequent establishment of a Sámi Parliament in Finnmark in 1989, while the Sámi Language Act, passed in 1992, recognized Northern Sámi as its official regional language. The Sámi Language Act saw the formal promotion of the language within the Sámi Parliament, in courts of law, and at all levels of education. In addition, a separate Sámi curriculum was introduced in Finnmark in 1997, and in 2000, the Sámi Parliament took responsibility for some aspects of the Sámi school system, previously controlled by the central Norwegian Government. Both these latter developments, along with the passing of the Finnmark Act (2005), have further entrenched regional autonomy and indigenous control for Sámi in the area (Huss 2017).

3.2. Established (national or regional) minorities

The category of 'indigenous peoples', however, does not extend in definitional terms to groups such as the Welsh, Catalans, and Basques, who have a similar historical association with a particular territory, but who do not share the characteristics usually associated with indigenous peoples (Darquennes 2013; Eriksen 2010). In this regard, Churchill's (1986) distinction between 'indigenous' and 'established' minorities is a useful addendum here. Churchill argues that 'established' and 'indigenous' minorities are both minority groups that have been long established in their native countries. However, where indigenous peoples are characterized by a 'traditional' culture often regarded as being at odds with that of the majority group, established minorities are characterized by a lifestyle similar to the remainder of the national society, although sometimes falling behind in rate of evolution. As such, established minorities are more likely to be able to lay claim to a right to conserve their identity, and to back such claims with political might.

The notion of 'established minorities' also closely accords with the category of 'national minorities' in international law, and the increasing rights attributable to them in supranational legislation (see below). Thus, while analytically distinct, it should be reiterated that the positioning of national minorities and indigenous peoples *vis-à-vis* linguistic marginalization remains similar. The political philosopher Will Kymlicka (1995) makes these connections explicit when he argues that both national minorities and indigenous peoples can be usefully collocated as groups of people who retain a historical association with a particular territory, but who have been subject to conquest, colonization, and/or confederation in that territory. This has resulted, he argues, in their comparable linguistic, and broader social and political, minoritization over time.

3.3. Ethnic (urban) minorities

Unlike the preceding groups, which can be seen as having an historic claim to a particular territory, modern migrants comprise those who have immigrated to and settled subsequently in a new country. Usually such groups have come to be concentrated in urban areas, although this may not have been the case historically. Often described as 'ethnic minorities', these groups may retain elements of their culture, language, and traditions − sometimes over the course of a number of generations − in the new host society. However, their general aim historically has been to integrate into the host society and to be(come) accepted as full members thereof. As such, ethnic minorities have traditionally been distinguished from established/national minorities and indigenous peoples on the basis that their ethnic, cultural, and linguistic distinctiveness is manifested primarily in the private domain, and is thus not inconsistent with their institutional integration into the nation-state (Kymlicka 1995). Accordingly, ethnic minorities do not usually seek separate, self-governing status within the nation-state, as is typically demanded by national minorities and indigenous peoples. Rather, they argue for a more *plural* and *inclusive* conception of national identity and culture, which recognizes their contribution to and influence on the historical and contemporary development of the host nation-state. In so doing, however, such claims may also potentially come to include greater public and/or institutional recognition of their distinctive cultural, religious, and/or linguistic backgrounds − i.e. a more *reciprocal* relationship between ethnic minorities and host nation-states.

The position of ethnic minorities in modern nation-states has become even more salient over the last decade or so due to the expansion and intensification of such migration (and more recently, transmigration) patterns. The British sociologist Steven Vertovec (2007) has described these processes as the advent of 'superdiversity'. As we shall see in the discussion of related issues in international law in the next section, the language rights of indigenous peoples and national minorities on the one hand, and ethnic minorities on the other, are though seldom articulated in either academic or legal contexts. As Extra and Gorter (2001: 3−4) noted over 15 years ago, in an observation that still holds:

> Despite the possibilities and challenges of comparing the status of regional minority and immigrant minority languages, amazingly few connections have been made in the sociolinguistic, educational and political domain. […] Contacts between researchers and policy makers working with different types of minority groups are still scarce. […] Overall, we see disjointed research paradigms and circles of researchers which have very little or no contact, although they could learn a lot from each other […].

The subsequent rise of superdiversity on the one hand, along with a widespread flight from multiculturalism as public policy in many western states (Modood 2013) over the last 20 years or so on the other, adds even further urgency to addressing this ongoing academic and policy disjunction between the language rights of national and ethnic minorities. This is particularly important, given the increasingly parlous position of ethnic minorities in many states today.

4. Linguistic minorities in international law

Meanwhile, the discussion of language rights for linguistic minorities in international law is similarly bifurcated. Although treated legally as distinct categories of people, the language rights of national minorities and indigenous peoples have increasingly become a focus of recent legislation, albeit after a long period of neglect (de Varennes 1996; Xanthaki 2007). However, the language rights of ethnic minorities remain barely addressed (Skutnabb-Kangas and May 2017). This ongoing bifurcation in international law has implications for both the legal definition of what constitutes a linguistic minority, and the rights attributable to them, as well as related regional, national, and supranational language policy developments. Before turning in more detail to the legal implications of the ongoing distinctions made among linguistic minorities with respect to language rights, I first briefly outline the broader historical evolution of these minority language rights in international law, such as it is.

4.1. The origins of language rights

While there has been a general antipathy to the recognition of minority language rights in the post-Second World War period, for reasons that I explain shortly, this was not always the case. In the 19th century, for example, treaties were often employed for the protection of minority groups, initially on the basis of religion, and later on the grounds of nationality − including within the latter language and culture (Thornberry 1991). These practices culminated in the general organization of the League of Nations, established in the wake of the First World War. The League endorsed a range of bilateral treaties aimed at securing special political status for national minority groups within Europe, in what came to be known as the Minority Protection Scheme (MPS). The MPS − overseen by its Permanent Court of International Justice (PCIJ) − was primarily concerned with the protection of 'displaced' minorities in other nation-states, the result in turn of the reorganization of European state boundaries after the First World War. They included two principal types of measures: 1) individuals belonging to linguistic minorities, amongst others, would be placed on an equal footing with other nationals of the state; 2) the means of preserving the national characteristics of minorities, including their language(s), would be ensured (de Varennes and Kuzborska 2019).

In the most prominent legal ruling on these provisions − the *Advisory Opinion on Minority Rights in Albania* (1935) − the PCIJ stated that these two requirements were inseparable. It concluded that "there would be no true equality between a majority and a minority if the latter were deprived of its own institutions and were consequently compelled to renounce that which constitutes the very essence of its being a minority" (Thornberry 1991: 399–403). On the basis of this judgment, linguistic minorities were confirmed in their right to establish private schools and institutions. However, where numbers warranted, another key principle in international law with respect to minority protection, public funding of minority language-medium schools, was also advanced. In respect to this and other similar decisions, linguistic minorities were defined purely on a numerical basis − that is, as constituting less than 50 % of the population. That said, freedom of choice as to membership in a minority also seemed to permeate the MPS.

4.2. Contesting language rights

Subsequent developments in international law, most notably the United Nations' 1948 *Declaration of Human Rights* (UNDHR), were to rapidly supersede these treaties and the principles upon which they were based. Minority language and education rights were largely subsumed within the broader definition of human rights adopted by the UNDHR. Human rights were thought, in themselves, to provide sufficient protection for minorities. Accordingly, post-Second World War, no additional rights were deemed necessary for minority group members.

Some 70 years on, this view still holds wide sway (May 2011). And yet, what is also increasingly apparent over this period is a clear pattern highlighting how a generalist human rights framework has consistently disadvantaged minorities, who have been left subject to the majoritarian decision-making processes of nation-states. As Kymlicka (1995: 5) observes, the result has been to render minorities "vulnerable to significant injustice at the hands of the majority, and to exacerbate ethnocultural conflict". As a consequence, the limits of a generalist human rights framework for minorities have led to echoes, albeit weak ones, of the earlier principles of minority protection in some key legislation, particularly with respect to language and education. The most notable of these has perhaps been Article 27 of the 1966 *International Covenant on Civil and Political Rights* (ICCPR), which imposes a *negative* duty on nation-states with respect to the protection of the languages and cultures of minority groups:

> The [UN] Sub-Commission preferred that 'persons belonging to minorities' should replace 'minorities' because minorities were not subjects of law and 'persons belonging to minorities' could easily be defined in legal terms. On the other hand, it was decided to include 'in community with other members of their group' after 'shall not be denied' in order to recognise group identity in some form (Thornberry 1991: 149).

The tension evident here between individual and group ascription is reflected in the question of who exactly can claim rights under Article 27. This question has been tackled on two fronts. First, following the precedent set by the earlier minority treaties, 'minorities' in Article 27 have come to be defined strictly in numerical terms. A minority is defined as a group of people who share in common a culture, a religion, and/or a language, and who constitute less than 50 percent of a *state's* population. Thus, a minority may be numerically dominant in a particular province − a *regional majority*, in effect, as, for example, the Québécois in Québec and the Catalans in Catalonia − but may still be classified as a minority within the nation-state.

Second, any person may claim to be a member of a linguistic minority group on the basis of self-ascription. However, to benefit from Article 27, they must also demonstrate that some *concrete* tie exists between themselves and the minority group. In relation to a minority language, this would require a real and objective tie with that language. It would not be sufficient, for example, to be a member of a minority ethnic group that is known to speak a particular language if the individual does not speak that language. Nor are particular languages, and the rights associated with them, tied to specific ethnic groups, since more than one ethnic group may speak the same language. Determining that an individual belongs to a particular linguistic minority is thus not an issue of establishing some type of legal or political category; it is principally an objective deter-

mination based on some concrete link between an individual and a linguistic community (de Varennes 1996).

4.3. In search of a definition

The definition of what constitutes a 'linguistic minority' for the purposes of Article 27 has continued to be contested in debates about the legitimacy of any minority group-based claims to language rights within international law up to the present day (de Varennes and Kuzborska 2019; May 2011). Perhaps the most widely agreed-upon definition of a linguistic minority over this period has been that of the Special Rapporteur, Francesco Capotorti (1991: 5), for the Sub-Commission on Prevention of Discrimination and Protection of Minorities, who in discussing the implications of the ICCPR, suggested a minority for the purposes of Article 27 could be defined as follows: "[a] group numerically inferior to the rest of the population of a State, in a non-dominant position, whose members − being nationals of the State − possess ethnic, religious or linguistic characteristics differing from those of the rest of the population and show, if only implicitly, a sense of solidarity, directed towards preserving their culture, traditions, religion or language".

Capotorti's definition has been the one most often referred to in subsequent discussions of language rights, including in recent UN documents (see, for example, "Minorities under international law" online at http://www.ohchr.org/EN/Issues/Minorities/Pages/internationallaw.aspx). However, it has also increasingly been contested on the basis that it only affords these rights to 'nationals', or citizens. This allows scope for states to disavow minority rights − or any rights, for that matter − to non-citizen minorities, an increasingly likely possibility, given the exponential rise of stateless peoples in the current era. The ethnic cleansing of the minority Muslim Rohingya by Myanmar in 2017–2018 starkly illustrates the potential consequences of such state action, although there has been precedent for this elsewhere, particularly in relation to the longstanding denial of rights to indigenous peoples by numerous states (Tully 1995).

Accordingly, the UN Human Rights Committee has since opted for a far less restrictive interpretation than the Capotorti definition, arguing that any linguistic, religious, or ethnic minority in a state is entitled to claim the minimum obligations and rights guaranteed by Article 27 of the ICCPR, regardless of status, dominance, period of residency, etc. (de Varennes and Kuzborska 2019). As de Varennes and Kuzborska also note, this position is reinforced by Article 31(1) of the *Vienna Convention on the Law of Treaties,* which specifies that a treaty must be interpreted in good faith, in accordance with the ordinary meaning of its terms, in their context, and in light of its object and purpose − in this instance, the definition of a linguistic minority. This understanding also extends to the protection of regional majorities who continue to be deemed a minority within the wider nation-state, as discussed previously, for the purposes of Article 27. This latter position is most clearly reflected in the case of Ballantyne, Davidson, and McIntyre v. Canada (1989), brought before the United Nations Human Rights Committee (UNHRC; see Communications Nos. 359/1989 and 385/1989, U.N. Doc. CCPR/C/47/D/359/1989 and 385/1989/Rev.1 1993). In the case, the litigants argued that, as English speakers in Québec, they were disadvantaged by the region's French language laws, and that as such, they should be treated as a linguistic minority under Article 27 of the ICCPR, with the

related right to English-medium education. The UNHRC court ruling, however, denied their claim, on the basis that "English-speaking citizens *of Canada* cannot be considered a linguistic minority". This is also in broad accord with the requirement in Article 27 for states to act against third parties, such as majority citizens, who intend to infringe, or have already infringed, upon this minority right (I thank one of the anonymous reviewers for drawing this point to my attention).

4.4. Reclaiming minority language rights

These ongoing definitional debates notwithstanding, it is clear that, over the last 30 years or so, there has been a growing, albeit still limited, accommodation of minority language rights in international law, particularly for national minorities and indigenous peoples. I briefly highlight here three key legislative developments as examples of both the potential affordances, and still considerable constraints, attendant to such developments. The first two legislative developments, the Council of Europe's *Framework Convention for the Protection of National Minorities* (1994) and the *European Charter for Regional or Minority Languages* (1992), relate to national minorities (including regional majorities). The third, the *United Nations Declaration on the Rights of Indigenous Peoples* (2007), is indicative of comparable developments for indigenous peoples (see also Chapter 59).

In November 1994, the *Framework Convention for the Protection of National Minorities* was first adopted by the Council of Europe and came into force in February 1998. The *Framework Convention* allows for a wide range of language rights for national minorities, including language and education rights. It also asserts at a more general level that contributing states should "promote the conditions necessary for persons belonging to national minorities to maintain and develop their culture, and to preserve the essential elements of their identity, namely their religion, language, traditions and cultural heritage" (Art. 2.1). That said, the specific provisions for language and education remain sufficiently qualified for most states to avoid them, if they so choose (Skutnabb-Kangas and May 2017).

A similar pattern of general aspiration, and specific qualification, with respect to language rights can be found in the *European Charter for Regional or Minority Languages*. The *Charter*, which was adopted by the Council of Europe in June 1992 and came into force in March 1998, aims to promote greater formal recognition of regional and minority languages within the Pan-European Community, and in so doing, to extend the provision of language education therein. It provides a sliding scale of educational provision for national and regional minority languages (but not for immigrant languages) that ranges from a minimal entitlement for smaller groups − preschool provision only, for example − through to more generous rights for larger minority groups, such as primary/elementary and secondary school language education. Again, however, nation-states have discretion in what they provide, on the basis of both local considerations and the size of the group concerned. European nation-states also retain considerable scope and flexibility over which articles of the charter they actually choose to accept in the first place. In this respect, they are only required to accede to 35 out of 68 articles, although 3 of the 35 articles must refer to education (Council of Europe 2008).

The tensions between the development of international legal treaties on minority rights, and their subsequent effective implementation by nation-states, are also clearly

reflected in the 2007 *United Nations Declaration on the Rights of Indigenous Peoples* (UNDRIP). The UNDRIP was formulated over a 25-year period. This included the development over more than 10 years of the 1993 *Draft Declaration by the Working Group on Indigenous Populations* (WGIP), which was in turn a part of the United Nation's Sub-Commission on the Prevention of Discrimination and Protection of Minorities. The merits of the *Draft Declaration* were subsequently debated for nearly 15 years, with many UN member states raising substantive and repeated objections to its promotion of greater self-determination for indigenous peoples (see Xanthaki [2007] for a useful overview). Despite these objections, the final version of UNDRIP retained its strong assertion of indigenous rights, *including* specific language and education rights. Article 14.1 states, for example, that "Indigenous peoples have the right to establish and control their educational systems and institutions providing education in their own languages, in a manner appropriate to their cultural methods of teaching and learning". Even so, its effective implementation within individual nation-states remains moot.

5. Conclusion

Developments in international law, then, are at once both encouraging and disappointing. The principle of separate minority recognition in language, including education for national minorities and indigenous peoples, is increasingly accepted as a valid legal position. However, such recognition still remains largely dependent on the largesse of individual nation-states, particularly with respect to active state support of these minority languages, including the provision of minority language education. As a result, there are as yet no watertight legal guarantees for the recognition and funding of minority language and education rights for national minorities and indigenous peoples, although significant advances have clearly been made in recent years.

Even so, there *is* increasing recognition within international and national law that significant minorities within the nation-state have a *reasonable* expectation to some form of state support (Carens 2000; de Varennes 1996). In other words, it would be unreasonable for nation-states to be required to fund language and education services for all minorities, for example. However, it is increasingly accepted that, where a language is spoken by a significant number within the nation-state − 'where numbers warrant' − it would also be unreasonable *not* to provide some level of state services and activity in that language. The concepts of 'reasonableness' and 'where numbers warrant' present a number of key possibilities for linguistic minorities. For national minorities and indigenous peoples, the notion of reasonableness provides a basis for moving beyond the strict numerical definition of minorities within international law in order to recognize their longstanding association with particular territories. However, even on the basis of 'where numbers warrant' alone, language rights can still be accorded to national minorities and/or regional majority language speakers. Canada, for example, adopts the principle of 'where numbers warrant', in relation to French speakers outside of Québec, via the *Canadian Charter of Rights and Freedoms* (1982). A similar approach is adopted in Finland with respect to first language Swedish speakers living there. Swedish speakers can use their language in the public domain in those local municipalities where there are a sufficient number of Swedish speakers (currently, at least eight percent) for these municipalities to be deemed officially bilingual.

India provides perhaps the best example of this principle in operation, with over 200 language varieties spoken across 30 states and five Union territories. On the one hand, we have seen in India the long-standing promotion of English, and more recently Hindi, as the state's elite, pan-Indian, languages. On the other hand, there are 22 languages (excluding English) that are recognized in India as 'principal medium languages', including two tribal languages (Mohanty and Panda 2017). The division of India's states along largely linguistic grounds means that, as a result, local linguistic communities have control over their public schools and other educational institutions. In addition, the Constitution of India (Article 350A) directs every state, and every local authority within that state, to provide 'adequate' educational facilities for instruction in the first language of linguistic minorities, at least at the primary/elementary school level. As a result, over 30 minority languages are employed as the medium of instruction in elementary schools throughout India. More broadly, these language rights, and related educational provisions, provide a key foundation for India's wider social and political stability, at least to date.

De Varennes and Kuzborska (2019) describe this approach to language rights in terms of the 'principle of proportionality'. A proportionate response to minority language provision should be considered positively by states, they argue, because such provision can increase linguistic minority participation in the public realm, enhance access to key public services, improve educational outcomes for linguistic minority students (see also May 2014a), and lead to wider economic and employment opportunities. Given that linguistic minorities are increasingly unwilling to accept the ongoing marginalization of their languages and cultures as the 'price' of citizenship in modern nation-states, such recognition can actually enhance the social inclusion of linguistic minorities − along with wider social cohesion − within modern nation-states, rather than undermining it (May 2018).

More positively, if nation-states are reimagined in more pluralistic and inclusive ways, there is the potential for recognition of not only greater political democracy, but greater ethnocultural and ethnolinguistic democracy as well (May 2012). This is increasingly recognized as an appropriate state response to the language claims of indigenous peoples and national minorities. However, the notion of 'where numbers warrant' also provides a basis for active state accommodation and support of ethnic minority languages as well − *at least, potentially*. This is significant, given that ethnic minorities continue to be largely ignored by international treaties on minority rights, along with a related pathologizing of their languages and cultures, in the current retreat from multiculturalism as public policy.

In contrast, a more pluralistic approach to language rights recognizes and addresses the increasing salience of language diversity in our globalized, late modern world. Recent work on sociolinguistic superdiversity (e.g. Arnaut et al. 2016; Blommaert 2013; May 2014b) foregrounds the complexities of individual multilingual repertoires in contexts of migration and transmigration. However, the issues traversed in this chapter suggest that more work still needs to be done in seriously addressing group-based claims to language rights, and their articulations at the national, supranational, and sub-state levels, despite − or, perhaps, more accurately, precisely because of − the inevitable complexities attendant to such claims.

6. References

Anderson, Benedict
2006 *Imagined communities: Rreflections on the origin and spread of nationalism*, revised edn. London: Verso.

Arnaut, Karel, Jan Blommaert, Ben Rampton & Massimiliano Spotti
2016 *Language and superdiversity.* New York: Routledge.

Bauman, Richard & Charles L. Briggs
2003 *Voices of modernity: Language ideologies and the politics of inequality.* Cambridge, UK: Cambridge University Press.

Blommaert, Jan
2013 *Ethnography, superdiversity and linguistic landscapes.* Bristol, UK: Multilingual Matters.

Brasil
1996 *Constituição da República Federativa do Brasil (CF/88)*. São Paulo: Editora Revista dos Tribunais.

Capotorti, Francesco
1991 *Study on the rights of persons belonging to ethnic, religious and linguistic minorities*. New York: United Nations, Dept. of Public Information.

Carens, Joseph
2000 *Culture, citizenship and community: A contextual exploration of justice as evenhandedness*. Oxford, UK: Oxford University Press.

Churchill, Stacy
1986 *The education of linguistic and cultural minorities in the OECD countries*. Clevedon, UK: Multilingual Matters.

Council of Europe
2008 *The European Charter for Regional or Minority Languages: Legal challenges and opportunities.* Strasbourg: Council of Europe Publishing.

Darquennes, Jeroen
2013 Language policy and planning in indigenous language minority settings in the EU. *Revue française de linguistique appliquée* XVIII(2). 103−119.

de Varennes, Fernand
1996 *Language, minorities and human rights*. The Hague: Kluwer Law International.

de Varennes, Fernand & Elžbieta Kuzborska
2017 International law and language minority education. In Teresa McCarty & Stephen May (eds.), *Encyclopedia of language and education, 3rd edn: Language policy and political issues in education*, 143–154. New York Springer.

de Varennes, Fernand & Elžbieta Kuzborska
2019 Minority language rights and standards: definitions and applications at the supranational level. In Gabriele Hogan-Brun & Bernadette O'Rourke (eds.), *Handbook on minority languages and communities*. London: Palgrave Macmillan.

Eriksen, Thomas
2010 *Ethnicity and nationalism: Anthropological perspectives*, 3rd edn. London: Pluto Press.

Extra, Guus & Durk Gorter (eds.)
2001 *The other languages of Europe: Demographic, sociolinguistic and educational perspectives*. Clevedon: Multilingual Matters.

Gellner, Ernest
1983 *Nations and nationalism: New perspectives on the past*. Oxford: Basil Blackwell.

Hobsbawm, Eric
1990 *Nations and nationalism since 1780*. Cambridge: Cambridge University Press.

Huss, Leena
2017 Indigenous and minority languages of Scandinavia and Finland. In Teresa McCarty & Stephen May (eds.), *Encyclopedia of language and education: Language policy and political issues in education*, 3rd edn., 367–381. New York: Springer.
Ives, Peter
2010 Cosmopolitanism and global English: Language politics in globalisation debates. *Political Studies* 58(3). 516–535.
Kymlicka, Will
1995 *Multicultural citizenship: A liberal theory of minority rights*. Oxford, UK: Clarendon Press.
Makalela, Leketi
2017 Language policy and education in Southern Africa. In Teresa McCarty & Stephen May (eds.), *Encyclopedia of language and education: Language policy and political issues in education*, 3rd edn., 519–530. New York: Springer.
May, Stephen
2011 Language rights: The "Cinderella" human right. *Journal of Human Rights* 10(3). 265–289.
May, Stephen
2012 *Language and minority rights: Ethnicity, nationalism and the politics of language*, 2nd edn. New York, NY: Routledge.
May, Stephen
2014a Contesting public monolingualism and diglossia: Rethinking political theory and language policy for a multilingual world. *Language Policy* 13(4). 371–393.
May, Stephen (ed.)
2014b *The multilingual turn: Implications for SLA, TESOL and bilingual education*. New York, NY: Routledge.
May, Stephen
2016 Language, imperialism and the modern nation-state system: Implications for language rights. In Ofelia García & Nelson Flores (eds.), *Oxford handbook on language and society*, 35–53. New York: Oxford University Press.
May, Stephen
2018 Language rights and language repression. In James Tollefson & Miguel Perez Milans (eds.), *Oxford handbook of language policy and planning*, 236–253. New York: Oxford University Press.
Minority Rights Group
1997 *World directory of minorities*, 2nd edn. London: Longman.
Modood, Tariq
2013 *Multiculturalism*, 2nd edn. Cambridge: Polity Press.
Mohanty, Ajit & Minati Panda
2017 Language policy and education in the Indian Subcontinent. In Teresa McCarty & Stephen May (eds.), *Encyclopedia of language and education: Language policy and political issues in education*, 3rd edn., 507–518. New York: Springer.
Poirier, Christian
2017 Brazil's national indigenous movement: Resolute in times of crisis. https://www.opendemocracy.net/protest/brazil-national-indigenous-movement/. (accessed on 13 August 2018.)
Skuttnab-Kangas, Tove & Stephen May
2017 Linguistic human rights in education. In Teresa McCarty & Stephen May (eds.), *Encyclopedia of language and education: Language policy and political issues in education*, 3rd edn., 125–142. New York: Springer.

Thornberry, Patrick
1991 *International law and the rights of minorities*. Oxford, UK: Clarendon Press.
Tully, James
1995 *Strange multiplicity: Constitutionalism in an age of diversity.* Cambridge: Cambridge University Press.
Vertovec, Steven
2007 Super-diversity and its implications. *Ethnic and Racial Studies* 30(6). 1024–1054.
Xanthaki, Alexandra
2007 *Indigenous rights and United Nations standards*. Cambridge: Cambridge University Press.

Stephen May, Auckland (New Zealand)

32. Language standardization

1. Introduction: Language standardization as an aspect of language contact
2. The Haugen model of standardization
3. Modelling dialect/standard constellations
4. Contemporary standardization of autochthonous minority languages
5. References

1. Introduction: language standardization as an aspect of language contact

Responding to a basic need in the human species for both social cooperation and social differentiation, natural languages are characterized by being not only sufficiently homogeneous to allow for transfer of *linguistic* meaning, but also sufficiently heterogeneous to allow for transfer of *social* meaning. The conflicting need for both sameness and difference underlies the dynamism of norm and variation (and change) in language, and feeds the negotiations of linguistic and social meaning in any kind of situation where people and their languages meet, from the microsocial interactions of individual people in their everyday lives, to the macrosocial contacts of communities of people throughout centuries. In social-psychological terms, these negotiations are processes of *accommodation* that take the form of either *convergence* or *divergence*, depending on the contextual circumstances (Giles and Coupland 1991: 60–93), and enhance either sameness or difference in language use. The notions of convergence and divergence, and their social-psychological foundation, are central to the sociolinguistic endeavor of describing and explaining processes of change in situations of contact between linguistic varieties (as evidenced by many publication titles, e.g. Auer, Hinskens, and Kerswill 2005; Braunmüller and House 2009). This holds true even for the process of change known as standardization – and probably for language standardization in general, although I shall delimit exemplification to the European arena.

https://doi.org/10.1515/9783110435351-032

Reflecting the conflicting needs for both homogeneity and heterogeneity as constitutive aspects of a speech community and its linguistic means of communication, European scholarly and language-political thinking features a centuries-old ideological tension between *rationalistic* and *romantic* approaches to the issue of language standardization (instructively discussed in Geeraerts [2003], and also conceptualized in sociolinguistics as discourses of anonymity and authenticity, e.g. Woolard [2008]). In the rationalistic tradition, the development of a common, unitary language is conceived of as a prerequisite for the kind of mutual understanding and free communication that is needed to expand and consolidate markets, develop the public sphere, and strengthen the nation. The standard language is seen and propagated as a socially neutral means of communication, available for everybody to appropriate as a key to education and to participation in society at large, while nonstandard varieties (the dialects) are construed as major barriers to social emancipation. In contrast, the romantic tradition rejects the pleaded social neutrality of the standard as an ideological phantasy, and construes it instead as an instrument of social oppression and exclusion. The rationalistic approach has its focus on the *linguistic* meaning-making function of language, and acclaims the advantages of convergence and uniformity; the romantic approach underscores the expressive *social* meaning-making function of language, and praises the value of divergence and variety. These approaches will always be present in any speech community as an attitudinal and ideological tension, favoring either convergence at the expense of variety, or divergence at the expense of uniformity, depending on the concrete circumstances of language contact.

Language standardization is normally conceived of as convergence: it strengthens homogeneity at the expense of heterogeneity. *Linguistically*, standardization reduces variability in a community's use of language; *ideologically*, standardization strengthens a shared belief in the community that there is a 'best' language that all members should strive to acquire and master; *socially*, standardization expands (the requirement for) standard language use to more and more domains of societal life, and to more and more members of the society. Even if we may subscribe to the view that standardization basically reduces variability and creates uniformity, we should keep in mind that *divergence* is always part of the picture. The tension between convergence and divergence − in its linguistic as well as attitudinal appearance − will have to be reflected, somehow, in any endeavor to model standardization as an aspect of language contact. This chapter illustrates the issues as they appear in: (1) models of standardization as an historical process, (2) models of standard/dialect constellations throughout European history, and (3) contemporary endeavors to standardize autochthonous minority languages.

The point of models is to capture features that are common across a variety of situations. In this chapter, models are used to focus on the commonalities that characterize language standardization as an aspect of language contact, relying on the use of exemplary cases; these models are not used as frameworks for understanding a large number of different situations (overviews and discussions of standardization phenomena across families can be found for Germanic languages in Deumert and Vandenbussche [2003], for Slavic languages in Kempgen et al. [2014], and for Romance languages in Lebsanft and Tacke [2018]). The perspective taken in this chapter is predominantly a top-down one which focuses on the significance of authorities and institutions. A possible bottom-up perspective on standardization, understood as the emergence of new varieties resulting from accommodation and convergence in interaction, falls outside the scope of this chap-

ter (see Le Page and Tabouret-Keller [1985], Schneider [2007], and Trudgill [2011] for details).

2. The Haugen model of standardization

The classic model of language *standardization as process* was developed by Haugen (1966a, 1966b, 1987), and has been used in a number of comparative descriptions, most illustratively perhaps in Deumert and Vandenbussche (2003) and Darquennes and Vandenbussche (2015). Together, these two volumes cover a broad gamut of standardization situations, as the former deals with major Germanic languages, and the latter with a variety of European minority languages. Haugen's model emerged from his analyses of Norwegian endeavors to break with Danish and to standardize Norwegian (Haugen 1966b; Jahr [2014] resumes and continues the analysis into present times). Norway was under the Danish throne from the end of the 14th century until 1814, was under the Swedish throne during the 19th century, and only obtained full independence in 1905. This background to Haugen's model might in itself (besides my own Dano-Norwegian identity) justify the notable predilection in this chapter for Danish as an exemplary case of standardization with origins in 16th century Renaissance/Reformation, and for Norwegian as an exemplary case of standardization with origins in 19th century European national liberation movements. Yet more importantly, I think these two cases deserve particular attention because the Danish and Norwegian speech communities have come out of the standardization processes that characterized the centuries of *modernity* and *nation-state building* in Europe as antipodes on the homogeneity-heterogeneity scale, and are entering the contact contexts of *late modernity* and *globalization* as the most and least homogenized, respectively, among European speech communities (Coupland and Kristiansen 2011; Kristiansen 2014). The histories of the two languages − which in many ways are one and the same shared history − provide the best of illustrations of how the convergence/divergence tension in standardization may result in very different outcomes.

This section is structured in accordance with Haugen's dividing of the standardization process into four stages: *selection*, *codification*, *elaboration*, and *implementation/acceptance*. These are best seen as dimensions that are more or less present and overlapping, in terms of both chronology and endeavors, as the standardization process proceeds.

2.1. Convergence and divergence in *selection*

In most situations, *selection* first involves divergence in the sense of *ideological* downgrading of an *exoglossic* standard (i.e. a transplanted, nonindigenous language associated with social elites and formal contexts of use). Next, selection involves divergence in the sense of delimiting the *linguistic* material on which to develop a new *endoglossic* standard based on indigenous language material. Basically, either a restricted or a broader gamut of variation can be selected. Because selection ideologically is about evaluative hierarchization of varieties, the broad-gamut choice presupposes an understanding of

linguistic variation and change as divergence from an earlier common source in terms of 'regular sound laws' instead of 'chaotic decline'. That understanding did not develop in the discipline of linguistics before the beginning of the 19th century, and has never enjoyed much broad appeal in the general public. Rather, whenever human communities develop language standards, these tend to fossilize linguistically, as they are ideologically considered the 'correct' and 'best' way with language. Subsequent changes are understood as 'decay', a morally reproachable break from a 'classic' and 'paradisiacal' linguistic past.

As this was the inherited way of thinking about language variation and change at the time of the Renaissance/Reformation (the classic languages, Latin in particular, had been unquestionably 'best' for centuries), and in the centuries to follow, the selection of an endoglossic basis had to be a search for the one 'correct' and 'best' variety among the chaos of 'corrupted' and 'vulgar' varieties. Not surprisingly, the grammarians of the time focused on the variety of the socially 'best' people: the people in power. In the kingdom of Denmark-Norway of the 16th–18th century, these were the social and cultural elites in the capital city of Copenhagen.

By the 19th century, the conditions for selection of a norm basis had changed with the realization that the patchwork of varieties which make up the Indo-European language family results from regular sound changes in different directions away from of a common ancestor. This made it possible to approach the selection issue in a new way. At the middle of the century, Ivar Aasen travelled around large parts of Norway and registered a wide gamut of dialectal variation, on the basis of which he constructed − applying the genetic-comparative approach to language variation − a (somewhat etymologizing) standard for writing Norwegian, intended to replace the Danish standard (Aasen 1848, 1850). The approach furthermore allowed Aasen to search for the 'genuine' − and in that sense 'best' − language among common people living in the countryside far away from the capital city: they spoke dialects which could be shown to have developed by 'sound laws' directly from Old Norwegian; they were the saviors of the Norwegian language. In contrast, the language of the social and cultural elites in the capital city (Christiania) was 'bad' language, 'corrupted' by Danish influence (Walton [1987] provides details on Aasen; also note that the Danish name 'Christiania' was replaced by the original Norwegian name 'Oslo' in politically independent Norway in 1924). A parallel standardization effort selected the language of the 'best' people − i.e. Christiania upper-class speech − as the basis for a gradual recodification of the written language used up to that point (which was now felt to be Danish). As a result of these two divergent selections of a norm basis, there are two standards for writing in Norwegian: *Nynorsk* and *Bokmål*.

2.2. Convergence and divergence in *codification*

Codification is the aspect of standardization that is concerned with stipulations, in the form of dictionaries and grammars, regarding how to spell and inflect words in writing (recommendations regarding how to pronounce words are also often included). Long-standing contact with exoglossic standards is reflected in the codification of European standards in many ways.

2.2.1. Script and orthography

The longstanding contact with Latin has had the consequence that the *script* used in all western and central European languages is the *Latin alphabet*; prior to this contact, the script used in the Nordic countries was the *runic futhark*. More or less conspicuous divergence is manifest in the use of diacritics added to letters and/or extra letters added to the alphabet (e.g. <æ, ø, å> in Danish and Norwegian). Besides allowing for a better rendering of the speech sounds for which the Latin alphabet lacks letters, such divergence may signal some kind of social distinctiveness, often national identity (Jaffe et al. 2012; Sebba 2007).

Examples of conflict between exoglossic-rooted and homegrown *orthographic* codification are legion in most languages. Such conflicts are often highly illustrative of the role played by social meaning and ideology in orthographic codification. A telling example is the former Danish practice of writing nouns with an initial capital letter, as in German, a practice which was banned by an orthographic reform adopted by Parliament in 1948. The reform also stipulated that the /ɔ/ sound should no longer be written as <aa>, but rather as <å>, in accordance with the Swedish and Norwegian orthographies. Both these changes had been on the agenda of language scholars and reformers for more than a century, in vain. Their easy adoption in 1948 can only be understood against the backdrop of a postwar ideological reorientation, which implied divergence from all things German and convergence within Scandinavia.

Most of the orthographic problems that preoccupy codifiers today follow from massive contact with English. Should newly imported words be incorporated with their English spelling, or should they be adapted to national spelling conventions? Rooted as they are in a relatively recent striving for national independence, the codifiers of Norwegian (both the *Nynorsk* and *Bokmål* norms) are more inclined to 'nationalization', allowing e.g. *gaid*, *sørvis*, and *teip* alongside *guide*, *service*, and *tape*, whereas the English spelling is the sole variants accepted in official Danish orthography.

Double forms like *guide* and *gaid* are purely orthographic variants; they do not reflect differences in speech. More importantly, both the *Nynorsk* and *Bokmål* orthographic norms are codified according to principles that grant writers the possibility of spelling (and inflecting) words in accordance with the basic features of their own respective dialects. At the same time, both norms codify variation in a way that makes them partially overlap, which allows people to favor either linguistic convergence or divergence in their writing.

2.2.2. Grammar and lexicon

The consequences of contact with Latin have been particularly weighty and burdensome in regard to *grammatical* codification. The tradition of conceptualizing grammar in any language in accordance with the categories of Latin grammar has been strong in western linguistics, and efforts to break away from the Latin model are relatively recent. In the *lexical* field, European languages share much material of Greek-Latin origin, in particular in the domains of abstract and science-related vocabulary, and common codification of such terminology may be said to represent convergence among languages-in-contact

by ideologically adhering to internationalism as a principle of language policy. However, the eagerness to comply with convergence of this kind varies among communities. Some will prefer to diverge from Greek-Latin, and create new words (neologisms) based on their own linguistic material − ideologically in the name of, say, democratic and pedagogical principles. Iceland is commonly mentioned as a community where much effort is being put into building words with Icelandic building blocks − an approach known as *linguistic purism* (Thomas 1991). Enjoyable and oft-cited examples are *tölva*, which combines *tala* ('number') and *völva* ('prophetess') to mean 'computer', and the noun *þyrla*, which is derived from the verb *þyrla* ('to whirl') to mean 'helicopter'. If the neologisms do not succeed in replacing the international words, the result of such divergence may be the coexistence of synonyms, and hence increased lexical variation. For instance, the neologism *folkestyre* ('government by the people') is commonly used in Danish in parallel with *demokrati* ('democracy').

Codifiers of today are, first and foremost, preoccupied with issues of purism versus *laissez-faire* in relation to English. Contact with English, as the current exoglossic standard of globalization, has reinvigorated the tension between rationalism and romanticism in debates about how to deal with the influx of Anglicisms. Should codification 'nationalize' (*gaid*) and create new words (*tölva*), or should codification leave Anglicisms in peace (*guide*, *computer*)? Danish codifiers adhere to the latter view, although rather inconsistently: some nouns are codified with the English plural *-s* (*drinks*, *joints*), some with *-s* as an alternative to the traditional Danish plural endings (*jobs*/*job*, *slogans*/*sloganer*), while still others do not allow plural *-s* (*guider*, *weekender*).

2.3. Convergence and divergence in *elaboration*

Elaboration is Haugen's term for efforts to develop expressive means: terminologies, styles, and genres. For languages that began their standardization history with the Renaissance/Reformation, the great elaboration advancements are linked to the 18th century and the many writing activities of the great spirits of the Age of Enlightenment. In Denmark-Norway, the leading figure in this regard was the Copenhagen professor Ludvig Holberg, who developed the use of Danish in a wide gamut of academic and other literary genres, including a series of satirical comedies in which he often ridiculed pompous ways of speaking with a dense use of words and expressions from the exoglossic standard languages of that time (French, German, and Latin). Holberg is often referred to as the creator of the Danish (written) language.

In the Norwegian national-liberation situation of the 19th and 20th centuries, it follows from their differing conceptions of the national nature of the former written standard that the *Nynorsk* and *Bokmål* camps must differ in their views on elaboration as a dimension of the standardization process. For the *Nynorsk* camp, stressing the Danishness of the texts of the common Dano-Norwegian past, it has always been a prerequisite to produce texts in the codified new norm, in more and more text genres. An elaboration task in that sense has never really existed for the other camp, as they see *Bokmål* as a continuation of the common textual tradition and heritage.

Supporting the development of 'plain language' in the public sector, at the expense of 'bureaucratic language', is a main aspect of contemporary elaboration efforts in many

countries. For example, the cooperation of the Nordic national language councils includes a regularly recurring conference on *klarsprog* ('clear language').

2.4. Convergence and divergence in *implementation/acceptance*

As indicated by the term *implementation/acceptance*, there are two opposing perspectives to be taken into account: implementation stands for actions to spread the standard norm 'from above' (the top-down perspective), while acceptance stands for converging reactions 'from below' (the bottom-up perspective). For standardization to succeed, efforts from above need to generate positive resonance from below. The criterion for success is whether the (selected, codified, and elaborated) norm is being spread to, and accepted by, more and more people in more and more societal domains. Typically, standardization of writing approaches the aim long before anything similar occurs with speech, if the latter ever happens. Influence from writing is commonly thought to be a standardizing factor in speech. However, in terms of spread throughout the population, evidence for this is not strong in the cases of Danish and Norwegian. It was only with the strong development of broadcast media from the 1960s onwards, TV in particular, that mediated language began to seriously influence the relative societal strength of standard and nonstandard speech − with very different results in the two countries (Kristiansen 2014).

In Denmark, no other option than standard (Copenhagen) speech was thinkable as the language of the new public sphere. Copenhagen speech became the normative target for young Danes all over the country, and most of them already spoke like Copenhageners by 1980, with traces of localness reduced to prosody. The general divergence in Danish society away from local ways with language, and general convergence towards Copenhagen ways, ideologically and linguistically, is amply documented in sociolinguistic research (Kristiansen 2009; Maegaard et al. 2013). Denmark indeed exemplifies successful implementation/acceptance of a standard language, even in speech.

In contrast, the Norwegian speech community at the onset of the TV era had not developed anything like the 'normal' Danish standard (whether in writing or speech). Instead, as a rapidly expanding and increasingly influential public sector of society, the media universe made dialectal diversity a 'natural' ingredient of people's everyday lives. To the extent that Norwegians had previously believed that a 'best' language is needed for public communication to function, that belief was thoroughly undermined by their everyday media experience. Today, dialect is widely accepted in any context of use, including reading of prime time news on national TV and of the Liturgy in church (Sandøy 2011).

The Danish and Norwegian communities are antipodes with regard to what has come of the modern-era struggle between converging and diverging forces in the standard-dialects dimension, yet they are similar in having been successful at ousting exoglossic standards. Not least against the backdrop of centuries-long efforts and struggles to expel exoglossic standards, the contemporary openness to English in many domains of society causes concern and debate. It took four centuries from the Danish translation of the Bible before Danish became the common language of lectures at Copenhagen University in the mid-19th century. It took English less than half a century to become a common language of lectures in large segments of higher education in Denmark. To deal with this

situation, a notion of *parallel language use* has been developed, and largely embraced, in Nordic language policy, including in the domain of higher education (Harder 2009; Hultgren 2014).

3. Modeling dialect/standard constellations

The general issues involved in *standardization as process*, as described above in terms of Haugen's classic model, are basically the same in any case of standardization. However, standardization may also be described in terms of the different *dialect/standard constellations* that emerge in the process of standardization.

3.1. Auer's typology of standard/dialect scenarios

The model developed by Auer (2005, 2011) builds on the argument that "the emergence of endoglossic standard varieties out of the older vernaculars follows a general line of development with five consecutive stages (types of repertoires)" (Auer 2011: 485). The first stage corresponds to the pre-Renaissance/Reformation situation, with Latin as an exoglossic standard without any (realized) structural relationship to the vernaculars – graphically pictured by a dot (the exoglossic standard) floating above, and without any connection to, a large circle (the gamut of vernaculars/dialects). As the endoglossic standard becomes part of the picture, it too is represented by a dot above the circle of dialects, but now with dialects-standard connections (visualized as lines and arrows) that are realized and developed by users in ways that yield the next four stages. The model keeps these apart by taking the issue of medium into account (writing only versus both writing and speech), and by distinguishing between diglossic and diaglossic situations. In a diglossic situation, the relationship between standard and dialect is an either/or choice; in a diaglossic situation, it is a question of where usage falls on the standard-dialect continuum of variation. Over time, these ways with language (and the evaluative forces behind them) reduce the total amount of variation. The combined effects of horizontal convergence (between dialects) and vertical advergence (in which dialects approach the standard) – called *regional dialects* and *regional standards* – emerge as intermediate varieties, as traditional dialects gradually disappear. At the final stage, the graphic representation shows a small, low circle with a small distance to the dot above: the traditional dialects are lost, and the small circle of variation can be viewed as variation within the standard. Danish is arguably the best example there is of this last stage.

3.2. Kloss' *Abstand/Ausbau* framework

An influential conceptualization of the relationship between dialects and standards on a dialect continuum was developed by Kloss (1967) in terms of *Ausbau* languages and *Abstand* languages. In a dialect continuum, communication is not a problem between neighboring dialects, because linguistic differences are few; but as the number of differ-

ences accumulates across geographic distance, communication turns out to be difficult (or impossible) between widely separated dialects. Dialects which are mutually unintelligible as a result of linguistic distance are *Abstand* languages. If one dialect is standardized to serve as a common means of communication for speakers of a multitude of dialects, this is called an *Ausbau* language; this notion roughly corresponds to Haugen's codification and elaboration. There may be several *Ausbau* languages which *roof* different areas of the dialect continuum. The difference between roofing and subordinated varieties has also been discussed in terms of *autonomy* versus *heteronymy* (Chambers and Trudgill 1980: 10−14).

The Scandinavian dialect continuum is an oft-cited example. It stretches from the Danish-German border in the south to the Norwegian Arctic Ocean shores in the north, and from the Norwegian Atlantic Ocean shores in the west to the Finland-Swedish speakers around Helsinki in the east. Many speakers within this vast territory will have problems communicating with each other, unless they resort to the standard languages which roof their dialects, i.e. either the Danish standard, a Norwegian standard (*Nynorsk* or *Bokmål*), or a Swedish standard (Sweden-Swedish or Finland-Swedish). Even if dialects are in close contact and no communication problems exist, they may well be conceived of as belonging to different (standard) languages. This is because such conceptions typically reflect social and ideological realities, as opposed to linguistic realities. As a consequence, since social and ideological realities can (and do) change, dialects can (and do) change roofing. This is what happened when *Nynorsk* and *Bokmål* were elaborated, as part of the national liberation efforts, to roof the Norwegian dialects, which were no longer conceived of as belonging to Danish. An illustrative contemporary example is found in the ongoing elaboration of four standard languages − Bosnian, Croatian, Montenegrin, and Serbian − to roof different portions of the dialect continuum that formerly was roofed by the Serbo-Croatian standard. Such changes in roofing are likely to actuate linguistic divergence along political borders. This seems to have happened rapidly, and quite dramatically, after Scania (the southernmost part of Sweden) in 1658, as an outcome of war, passed from belonging under the Danish throne to belonging under the Swedish throne (Ohlsson 1978). This has also been found to be a present-day process in the German varieties spoken on both sides of the German-French border (Auer et al. 2015).

As a result of political dominance/independence and (re)drawing of borders throughout history, the world features a long list of so-called *pluricentric* languages, i.e. languages which are subjected to standardization efforts and used as official languages in different countries (e.g. German in Germany, Austria, and Switzerland). It can become an issue of debate and disagreement whether independent standardization should be encouraged or counteracted in such situations, e.g. the situation of Swedish in Finland (Östman and Mattfolk 2011), and the situation of Dutch in Belgium (Delarue 2013). Colonial languages (such as English, Spanish, French, and Portuguese) typically become pluricentric as new states engage in their own standardizing efforts, e.g. alternatives to the UK standard in the USA, Australia, New Zealand, South Africa, and so on.

4. Contemporary standardization of autochthonous minority languages

Throughout history, processes of language standardization have typically been instigated as an important ingredient of societal movements striving for national independence and unity. In general, this is still the situation and agenda today. At the same time, the situation has changed for the many autochthonous minority languages in Europe. The term *autochthonous* refers to the old language minorities that have lived in Europe for centuries (not to the many newly arrived language minorities from outside of Europe). These minorities and their languages share a long history of strong suppression, which in many cases have brought them to the edge of extinction, or even pushed them over that edge in some cases. Practically all speakers of these languages also speak the majority language of the state they live in. A major common problem is the decline in use of these languages by young people. With the general change in western thinking about indigenous peoples and cultures from the 1960s on, which in Europe is manifested in the *European Charter for Regional or Minority Languages* (adopted by the Council of Europe in 1992; see Darquennes [2011] for a critical presentation), standardization of autochthonous minority languages became an issue, not as a means to further anything like national independence, but mainly as a means to revitalize and maintain these languages. Many of them have in recent decades obtained various degrees of official recognition and status, typically yielding their speakers some language-related rights in limited geographic areas, especially in terms of education and communication with authorities. These new settings of language contact accentuate some of the crucial questions and dilemmas concerning the realizability and suitability of standardization as a means to an end.

Number of speakers is a critical aspect of this issue. The Sámi language is estimated to be mastered by some 25,000 people, and its recent history offers an example of standardization with promise for the future. The Sámi people inhabit a vast area spreading across the northern parts of Norway, Sweden, and Finland, and into the Russian Kola peninsula. Their language (which is related to, but not mutually intelligible, with Finnish) forms a dialect continuum, including several dialects that are mutually unintelligible. Six different orthographies exist. However, the great majority of Sámi-speaking people speak a variety known as North Sámi, which was selected for the codification of a common orthography. This orthography was developed by the Sámi Language Council in the 1970s, was subsequently accepted by the Nordic Sámi Conference, and has been used in Norway, Sweden, and Finland since 1979. This has allowed for the elaboration of a written language as an essential precondition for the legislative and institutional support that has developed for North Sámi in recent decades, including considerable progress in terms of implementation/acceptance in the essential domains of education and media.

An example of efforts to standardize in a quite different language contact situation is found in the Finnish dialects spoken on both sides of the Swedish-Finnish border river Torne. On the Swedish side, these dialects were (even more so than the Swedish dialects) in danger of extinction when efforts of maintenance and revitalization were initiated in the 1980s. The traditional Swedish name *Tornedalsfinska* (‘Finnish of the Torne Valley’) was replaced by the name *Meänkieli* (‘our language’), and in 1999, Meänkieli was grant-

ed official minority language status in Sweden (in Finland, Meänkieli is considered a dialect of Finnish). In spite of a considerable number of people who speak and/or understand the language (estimates vary a lot, from 30,000 to 70,000), as well as many initiatives and advancements in codification (dictionaries, grammars, and textbooks), elaboration (genres of literature), and implementation (local broadcasting), Meänkieli's closeness to Finnish seems to be a major problem when it comes to acceptance, or even recognition. The frequency of Standard Finnish words in standardized Meänkieli, in the place of Swedish loan words, often makes dialect speakers believe they are listening to Finland-Finnish. Reflecting 'rationalistic' rather than 'romantic' considerations, many young people in the area consider it more sensible to invest in learning Standard Finnish rather than Meänkieli (Lainio and Wande 2015).

The close relationship of Kven to Finland-Finnish is also a problem for the recent efforts to develop a standard to roof the Finnish dialects spoken in the northernmost part of Norway (all of which are mutually comprehensible with Finland-Finnish). Kven was officially recognized as a minority language in 2005 (estimates of numbers who speak/understand the language vary from 2,000 to 15,000). It is not that the codification of Kven was particularly Finnish-standard focused (it was instead inspired by the general Norwegian sense of dialect-related variability in writing), but rather that the elaboration of teaching materials has been sparse, such that implementation in terms of education is based on texts in standard Finnish. Acceptance is also a major problem, as traditional negative attitudes toward the language are still widespread among Kven speakers themselves. Many think of their language as 'broken Finnish', and hold negative or ambivalent attitudes toward the idea of developing a standard for Kven (Lane 2016).

Thus, the 'external divergence' issue (e.g. the linguistic and ideological relationships to the Finnish standard) emerges as a major problem for standardization in situations like that of Meänkieli and Kven (i.e. minority languages closely related to an official language in a neighboring country). At the same time, the 'internal convergence' issue (the linguistic and ideological relationships within the constructed dialects/standard constellation) also may well become a greater problem when the 'romantic' aim is limited to language defense (as against nation-state building). Thus, 'internal divergence' − resulting from a favoring of one dialect over and above others (whether real or imagined) − may inhibit general acceptance. Even if codifiers are aware of this danger and try to circumvent it, people may feel that it is a dialect other than their own that is reflected in writing. This is reported as a problem in the case of both Meänkieli and Kven. Similar *Abstand* and *Ausbau* issues − triggering bottom-up resistance to top-down implementation efforts − are reported to thwart standardization in many contact situations, including e.g. Ladin in Northern Italy (Videsott 2015), Mirandese in Portugal on the border with Spain (Martins and Santos 2015), and Galician in Spain on the border with Portugal (Ramallo and Rei-Doval 2015).

5. References

Aasen, Ivar
1848 *Det norske Folkesprogs Grammatik*. Kristiania [Oslo]: Feilberg & Landmark.
Aasen, Ivar
1850 *Ordbog over det norske Folkesprog*. Kristiania [Oslo]: Feilberg & Landmark.

Auer, Peter
2005 Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In Nicole Delbecque, Johan van der Auwera & Dirk Geeraerts (eds.), *Perspectives on variation*, 7–42. Berlin & New York: De Gruyter.
Auer, Peter
2011 Dialect vs. standard: A typology of scenarios in Europe. In Bernd Kortmann & Johan van der Auwera (eds.), *The languages and linguistics of Europe: A comprehensive guide*, 485–500. Berlin & New York: De Gruyter.
Auer, Peter, Frans Hinskens & Paul Kerswill (eds.)
2005 *Dialect change: The convergence and divergence of dialects in contemporary Europe*. Cambridge: Cambridge University Press.
Auer, Peter, Julia Breuninger, Dominique Huck & Martin Pfeiffer
2015 Auswirkungen der Staatsgrenze auf die Sprachsituation im Oberrheingebiet. In Roland Kehrein, Alfred Lameli & Stefan Rabanus (eds.), *Regionale Variation des Deutschen: Projekte und Perspektiven*, 323–347. Berlin & New York: De Gruyter.
Braunmüller, Kurt & Juliane House (eds.)
2009 *Convergence and divergence in language contact situations.* Amsterdam & Philadelphia: John Benjamins.
Chambers, J. K. & Peter Trudgill
1980 *Dialectology.* Cambridge: Cambridge University Press.
Coupland, Nikolas & Tore Kristiansen
2011 SLICE: Critical perspectives on language (de)standardization. In Tore Kristiansen & Nikolas Coupland (eds.), *Standard languages and language standards in a changing Europe*, 11–35. Oslo: Novus Press.
Darquennes, Jeroen
2011 Minorities, language politics and language planning in Europe. In Bernd Kortmann & Johan van der Auwera (eds.), *The languages and linguistics of Europe: A comprehensive guide*, 547–560. Berlin & New York: De Gruyter.
Darquennes, Jeroen & Wim Vandenbussche (eds.)
2015 *The standardisation of minority languages* (Sociolinguistica 29). Berlin: De Gruyter.
Delarue, Steven
2013 Teachers' Dutch in Flanders: The last guardians of the standard? In Tore Kristiansen & Stefan Grondelaers (eds.), *Language (de)standardisation in late modern Europe: Experimental studies*, 193–226. Oslo: Novus Press.
Deumert, Ana & Wim Vandenbussche (eds.)
2003 *Germanic standardization: Past to present*. Amsterdam & Philadelphia: John Benjamins.
Geeraerts, Dirk
2003 Cultural models of linguistic standardization. In René Dirven, Roslyn Frank & Martin Pütz (eds.), *Cognitive models in language and thought: Ideology, metaphors and meanings*, 25–68. Berlin & New York: De Gruyter.
Giles, Howard & Nikolas Coupland
1991 *Language: Contexts and consequences*. Milton Keynes: Open University Press.
Harder, Peter (ed.)
2009 *English in Denmark: Language policy, internationalization and university teaching* (Angles on the English-speaking world 9). Copenhagen: Museum Tusculanum Press.
Haugen, Einar
1966a Dialect, language, nation. *American anthropologist* 68(4). 922–935.
Haugen, Einar
1966b *Language conflict and language planning: The case of modern Norwegian*. Cambridge, MA: Harvard University Press.
Haugen, Einar
1987 *Blessings of Babel*. Berlin: De Gruyter.

Hultgren, Anna Kristina
2014 Whose parallellingualism? Overt and covert ideologies in Danish university language policies. *Multilingua* 33(1–2). 61–87.
Jaffe, Alexandra, Jannis Androutsopoulos, Mark Sebba & Sally Johnson (eds.)
2012 *Orthography as social action: Scripts, spelling, identity and power*. Berlin & New York: De Gruyter.
Jahr, Ernst Håkon
2014 *Language planning as a sociolinguistic experiment*. Edinburgh: Edinburgh University Press.
Kempgen, Sebastian, Peter Kosta, Tilman Berger & Karl Gutschmidt (eds.)
2014 *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte und ihrer Erforschung*. Berlin & New York: De Gruyter.
Kloss, Heinz
1967 *Abstand* languages and *Ausbau* languages. *Anthropological linguistics* 9(7). 29–41.
Kristiansen, Tore
2009 The macro-level social meanings of late-modern Danish accents. *Acta Linguistica Hafniensia* 41. 167–192.
Kristiansen, Tore
2014 Does mediated language influence immediate language? In Jannis Androutsopoulos (ed.), *Mediatization and sociolinguistic change*, 99–126. Berlin & New York: De Gruyter.
Lainio, Jarmo & Erling Wande
2015 Meänkieli today: To be or not to be standardised. *Sociolinguistica* 29. 121–140.
Lane, Pia
2016 Standardising Kven: Participation and the role of users. *Sociolinguistica* 30. 105–124.
Lebsanft, Franz & Felix Tacke (eds.)
2018 *Manual of standardization in the Romance languages*. Berlin & New York: De Gruyter.
Le Page, Robert B. & Andrée Tabouret-Keller
1985 *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
Maegaard, Marie, Torben Juel Jensen, Tore Kristiansen & Jens Normann Jørgensen
2013 Diffusion of language change: Accommodation to a moving target. *Journal of sociolinguistics* 17(1). 3–36.
Martins, Cristina & Isabel Santos
2015 Standardising Mirandese: Processes and challenges. *Sociolinguistica* 29. 141–156.
Ohlsson, Stig Örjan
1978 *Skånes språkliga försvenskning 1 Inledning: Om skriftspråket under Skånes övergångstid* (Lundastudier i nordisk språkvetenskap 30, Serie A). Lund: Walter Ekstrand Bokförlag.
Östman, Jan-Ola & Leila Mattfolk
2011 Ideologies of standardisation: Finland Swedish and Swedish-language Finland. In Tore Kristiansen & Nikolas Coupland (eds.), *Standard languages and language standards in a changing Europe*, 75–82. Oslo: Novus Press.
Ramallo, Fernando & Gabriel Rei-Doval
2015 The standardization of Galician. *Sociolinguistica* 29. 61–81.
Sandøy, Helge
2011 Language culture in Norway: A tradition of questioning standard language norms. In Tore Kristiansen & Nikolas Coupland (eds.), *Standard languages and language standards in a changing Europe*, 119–126. Oslo: Novus.
Schneider, Edgar W.
2007 *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.

Sebba, Mark
2007 *Spelling and society: The culture and politics of orthography around the world*. Cambridge: Cambridge University Press.
Thomas, George
1991 *Linguistic purism*. London & New York: Longman.
Trudgill, Peter
2011 *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.
Videsott, Paul
2015 Ladin Dolomitan: Die (vorerst) unterbrochene Standardisierung des Dolomitenladinischen. *Sociolinguistica* 29. 83–97.
Walton, Stephen J.
1987 *Farewell the spirit craven: Ivar Aasen and national romanticism*. Oslo: Det Norske Samlaget.
Woolard, Kathryn A.
2008 Language and identity choice in Catalonia: The interplay of contrasting ideologies of linguistic authority. In Kirsten Süselbeck, Ulrike Mühlschlegel & Peter Masson (eds.), *Lengua, nación e identidad: La regulación del plurilingüismo en España y América Latina*, 303–323. Frankfurt am Main: Vervuert & Madrid: Iberoamericana.

Tore Kristiansen, Copenhagen (Denmark)

33. Domains

1. The concept of 'domain'
2. Historical development
3. Domain loss
4. Critique of the domain concept
5. References

1. The concept of 'domain'

The concept of 'domain of language behavior' or 'domain of variance in language behavior' was developed in sociolinguistics from the late 1960s to the mid-1980s as a theoretical construct that could help explain the language choices of multilingual speakers in multilingual societies. The assumption has been that, in societies where everybody is at least bilingual and could use any of the two or more common languages freely, code-switching and language choice are not random, but rather patterned.

Like so many sociolinguistic concepts, 'domain' is often used in a rather loose way, without overly much consideration of the epistemological status of the term. In this way, it is often assumed that it is known beforehand which domains exist in a society, and that the task of the researcher is only to find out what goes on within these predefined domains. Some of the problems and quandaries arising from the application of the concept can best be discussed against the backdrop of the history of the concept.

https://doi.org/10.1515/9783110435351-033

2. Historical development

The origin of the concept of 'domain' is credited by Fishman (1986: 440) to German researchers of the 1930s who studied language maintenance and language shift in several groups of *Auslandsdeutsche*, that is, ethnic Germans outside Germany itself. Schmidt-Rohr (1933) was the first to describe 'dominance configurations' for language choice. For a particular setting (like German-speaking Switzerland with vernacular *Schwyzerdüütsch* and spoken and written *Schriftdeutsch*, the Alto Adige region with German and Italian, or the Eupen-Malmédy area in Belgium with French and German), he assumed nine elements of such 'dominance configurations': the family, the playground and 'the street' (understood here as a place where children also play), the school, the church, literature, newspapers, the army, the courts of justice, and public administration. In the case of the 'school' domain, a subdivision is made between the language of instruction, the subject of instruction (where foreign languages are a possibility), and the language used during breaks and conversations (Schmidt-Rohr 1933: 179).

From the mention of foreign languages as school subjects, and the choice of his examples in general, it becomes clear that Schmidt-Rohr did not necessarily consider (relatively) balanced multilingualism to be a condition for investigating language choice. His interest in the maintenance of German as a minority language, and his wish to distinguish 'harmless' from 'disastrous' multilingualism (Schmidt-Rohr 1933: 179) would make this less likely anyway, since he obviously tended to see maintenance in these situations as precarious. It is also clear that his elements of dominance configurations are applied equally to all his illustrative examples, which means that they are not construed on the basis of observations of which languages are chosen when, but rather on the basis of an *a priori* assumption of their relevance. Hence, the same set of elements can be applied to all types of configurations, which are then distinguished on the basis of the language use that corresponds to those elements (which were later called 'domains').

One could call this original domain concept the 'naïve domain concept', since it assumes that domains (the elements of dominance configuration), like the school, the family, and courts of justice, can be identified before any observation of language choice behavior. This is of course plausible in cases where institutions set up rules e.g. for the language of the courtroom or the classroom, but we also know that these rules are not always followed. A modern example is that in Swedish university classes nominally taught in English, a number of other languages (and not only Swedish) are often used by participants (Söderlundh 2014).

This impressionistic, or naïve, domain concept was applied by a number of researchers in the 1930s in Germany. Immediately after the war, this tradition, heavily tinged by a mother tongue ideology and a certain suspicion of multilingualism as potentially damaging for the individual (Hutton 1999), was not continued. It was only when Fishman (1964, 1965, 1972a) dusted off Schmidt-Rohr's writings that the use and spread of the domain concept gained momentum. The classic formulation was that "only one of the theoretically co-available languages will be chosen by particular classes of interlocutors on particular occasions" (Fishman 1965: 68), to which Fishman (1986: 437) later added "to discuss particular kinds of topics". It is important to emphasize that this discussion explicitly excludes situations of mere language contact, which had been at the core of the early German studies from the 1930s. This does not, however, mean that the domain

concept was not used in later language contact studies (Koufogiorgou 2008; Rindler Schjerve 1996).

Domains are neither given beforehand (Fishman 1986: 436), nor directly observable. They are, however, real for participants due to the mutual expectation consensus based on the role-relations of the participants, the (institutional and activity-related) settings, and the topics under discussion. For the analyst, they can be identified on the basis of observable language behavior related to the mentioned constituents of the interactional setting.

It has been emphasized not only by Fishman, but also by Saville-Troike (2003: 42), that there is no fixed set of domains that can be posited *a priori*. In Saville-Troike's discussion of the phenomenon, it is noted that the potential set of activities in any given situation is culturally specific, reflecting the cluster of purposes, role-relations, and settings that delimit the domain in question.

In Fishman's (1972a: 19) original definition, there was already reference to institutions in the form of "institutional contexts or socio-ecological cooccurrences". Since institutions can be interpreted in the same way as Schmidt-Rohr used 'army' or 'school', this paves the way for taking as a given the institutions present in a society with multilingual speakers, and looking for patterns of language choice within said institutions. Even several contributions in the early Bilingualism in the Barrio study (Findling 1971; Greenfield and Fishman 1971) make it clear that the domain approach − as promoted by Fishman, his associates, and the majority of his followers − gradually focused on identifying patterns of language choice within given domains, as opposed to identifying domains on the basis of observable language choice patterns within the given sociocultural (including institutional) context. This fits with dictionary definitions such as, "a domain is a combination of factors which are believed to influence choice of code (language, dialect or style) by speakers" (Trudgill 2003: 41).

3. Domain loss

While the concept of 'domain' has mostly been of interest to researchers within sociolinguistics (even before the field constituted itself under that name), the notion of 'domain loss' has gained recognition in the wider public, since it can be used in a discourse of language endangerment. Endangerment discourses most often concern 'small' languages in postcolonial settings that are on the verge of extinction. Yet endangerment discourses can also be about vigorous national languages used by large numbers of speakers in industrialized countries which appear to drift into a situation where the language is replaced by some other language in certain institutional contexts. An often discussed example is the increased use of English as a medium of instruction in secondary and tertiary education in several non-English speaking European countries, or the replacement of French and German as international languages of academic publication by English. It is not strange that the concept of domain reemerges in these discourses in the shape of domain loss: the original promoters of the domain concept were concerned with language maintenance and shift, and what is called domain loss can be described as a change in dominance configurations of language choice. There is a broad academic and nonacademic debate about domain loss, especially in Scandinavia (Harder 2009; Hult-

gren 2016; Hultgren, Gregersen, and Thøgersen 2014; Jarvad 2001), but also in other European countries (Oberreuther et al. 2012; Stickel 2010). De-standardization, in which a formal national variety is infiltrated and threatened by an informal 'substandard', can also be perceived as a type of domain loss for the formal variety (Kristiansen and Grondelaers 2013).

The criticism leveled at the use of the term 'domain loss' is mainly as follows. First, the societies involved are not multilingual societies with multilingual speakers, as presumed in original domain theory, rather there is at most some partial, unbalanced elite multilingualism for some parts of the population. Therefore, language choices are mostly externally determined (Salö 2017). Second, it has never been established that 'academia' or '(natural) science' actually are domains (Preisler 2005). Dominance configurations for language choice may differ considerably within academia (e.g. between teaching, research, publications, and administration), which makes it less likely that academia can be considered a single, consistent domain. The shift to English as a medium of academic instruction in many countries happened much later than, and independent of, changes in languages of academic publications. Publications in academia have for a very long time relied on international, not local, languages (although it may have shifted what is viewed as local versus as international at any given time). In Danish linguistics (to take an example from Haberland [2005]), there could have been at most a domain loss for French and German *vis-à-vis* English for international publications, not for Danish, which was rarely used beforehand anyway. Finally, domains as dominance configurations are not properties of languages, but of linguistic regimes of multilingual societies (Simonsen 2002, 2005). If domains are not properties of languages, it makes no sense to speak of languages as 'having' domains, and if a language cannot have domains, it cannot lose them either.

The phenomenon called 'domain loss' does exist, but it is doubtful whether it is adequately described as the loss of something that could be described as a 'domain'. Rather, it is more akin to the restructuring of language resources under external or internal pressure.

4. Critique of the domain concept

Criticism of the validity and applicability of the domain concept has been based on the realization that, while there is an awareness of the fact that the classical domain concept is a construct, i.e. not immediately observable but constructed in order to make sense of observable data, it is itself dependent on other constructs, like 'language' and 'multilingual society'.

First of all, it is assumed that language choice is based on the premise that proper usage dictates that *one* language is spoken at any time. The distinction made by Gafaranga (as in Torras and Gafaranga [2002]) between two types of language alternation (code-switching oriented to by participants, on the one hand, and bilingual medium where the speakers draw on two 'codes', but in effect treat them as a single code, or 'single medium' in Gafaranga's terminology, on the other) applies here, too: in many multilingual societies, speakers can sometimes chose between two or more languages in an exclusive manner, or alternatively choose to use several languages in the same situation,

even with the same speakers. Some of these cases might be described as metaphorical codeswitching (Blom and Gumperz 1972); however, that would contradict Fishman's claim that 'proper usage' favors the use of *one* language in any given situation. There are abundant counterexamples. Pedersen (2000) describes meetings in a Danish sports club in Southern Jutland, where proper usage requires Danish for the more formal, even ceremonial, portions of the meetings, and allows German for the more content-oriented discussions. Parkin (1974) reports conversations in markets in Nairobi, where languages are mixed: while the mixing is certainly not random, it does not, however, seem to follow any domain pattern. The 'market' could of course turn out to be a separate element of a dominance configuration, yet what is dominant is not a single language (Luo, Kikuyu, English, or Swahili), but rather a particular kind of bi/multilingual medium. In their analysis of urban multilingual ('metrolingual') practices, Pennycook and Otsuji (2015) introduce the concept of 'spatial repertoires'. In places like markets and kitchens, people need neither to converge on one common language, nor to assign languages to functional niches defined by domains, but rather to make use of all the sociolinguistic resources that are available in that particular space.

Tsitsipis' (1991) study of the use of a contracting language (Arvanitika) in a Greek context does not necessarily point at anything like domains as triggering language choice; it is instead the level of authoritativeness and abstraction of a statement that is decisive. On the other hand, it appears that in other studies (like Koufogiorgou [2008] on the temporary revitalization of Aromanian [Vlach] as a lingua franca in a Greek village when bilingual economic migrants from Albania arrived), the use of the domain concept makes sense.

One could object that the examples mentioned here are not set in multilingual societies proper, but are cases of language contact. Thus, they are not truly cases of "within-group multilingualism [where] general knowledge of mother tongue and other tongue may be ruled out as an operative variable since most individuals *could* communicate with each other quite easily in *either* of the available languages" (Fishman 1965: 67, emphasis in original). But this drastically restricts the number of cases relevant for domain analysis. Although Mackey (2000: 26) maintains that "a self-sufficient linguistic community has no reason to remain bilingual" − which would, if true, rule out within-group multilingualism, or at least relegate the term to referring exclusively to social configurations characterized by transiency in the sense of Mortensen (2017) − we have to ask whether the basic assumption of Fishman's, namely the very existence of balanced within-group multilingualism, is justified. Sociolinguistic situations are characterized by a certain fluidity, not only diachronically, in that they are experienced as moving from one state perceived as lasting to another one, but also synchronically, in that they often exhibit competing, coexisting normative orientations. If the domain concept should only be applicable to very fragile and precarious states of temporary balance, its usefulness would be severely restricted.

Another problem is whether language choice actually follows something that could be described as 'domains'. Ó Riagáin (1997) refers to the 1975 study of the Committee on Irish Language Attitude Research (CILAR), which set out to find the distribution of the use of Irish versus English according to domains. Domains were identified with institutional settings "such as home, work or public administration"; and "domain theory postulates that such variation is related [not to "persons" or "groups" but] to areas of 'activity'" (CILAR 1975: 230, as cited in Ó Riagáin 1997: 33). The problem was that

the use of Irish versus English did not seem to follow such predefined institutional frameworks. Again, one could object that this was to be expected, since Irish society is not a case of full within-group bilingualism: there are people who "*could* communicate with each other quite easily in *either* of the available languages", as Fishman (1965: 67) set out to say, but they do mingle and communicate extensively with people who could not; even traditional music events like *céilithe ceoil* are not exclusively Irish-language social events. In situations like these, metaphorical code-switching − i.e. what Fishman (1972b: 6; see also Blom and Gumperz 1972) calls "fleeting and non-reciprocal" switching − can occur as an expressive means, but is not domain-related.

If one allows for domains outside balanced within-group multilingualism, as in academia, one still has to construct the relevant domains carefully. Preisler (2005) has argued that 'science' is not a domain in Denmark, since it is not clear if 'science' as a whole can be considered a (or exactly one) cluster of 'congruent situations' which make up a domain according to Fishman's (1972a: 22) model. Science is not one "major cluster of interaction situations" (Fishman 1972a: 19), but at most a collection of such clusters. This has been underscored by Hamel (2008: 196), who distinguishes between at least three separate areas within 'science' for the analysis of language choice: *recherche* (research)*, diffusion* (publications), and *formation* (teaching). This tripartite distinction has been corroborated in empirical studies like Madsen (2008) for Denmark and Bretxa, Comajoan, and Vila i Moreno (2016) for Catalonia. Finally, the concept of language itself can be proplematized, since it is a construct in itself "whose 'objective' existence is limited to the lexicon and the grammar book" (Preisler 2005: 243). That does not mean that languages do not exist, but that the assumption of 'proper usage' requiring that one language is chosen is dependent on a clear delimitation between languages. This may be possible in written texts, where orthographic conventions (which are specific to a single language) index quite clearly which language we are dealing with, but this is less clear in oral interaction. In a radical concept almost doing away with languages altogether, and replacing them with features that are open to free combination (Jørgensen et al. 2011), the concept of language choice becomes superfluous, and with it, both the need and space for the concept of domain. On the other hand, the concept of 'translanguaging' (García and Li Wei 2014), which now covers the strategic use of several languages not just in pedagogical contexts, does not rely on reference to domains, since the choice of languages is not dependent on appropriateness in a given setting, but rather on suitability for the intended aims of language use.

In the literature, several suggestions have been made which could be taken as approaches either to replace the domain concept or to remove the need for its application. Such suggestions usually reinstate a factor that has been ruled out or marginalized as explanatory by earlier researchers. Candidates suggested by Ó Riagáin (1997) are group membership, as in Milroy's (1980) network analysis, or Bourdieu's (1991) notion of linguistic capital. This only underscores that the complex reality of variation in language behavior can be approached from several angles, and that analytical concepts may be exclusive *vis-à-vis* each other (since they are built on incompatible, divergent idealizations of a complex situation), but not *vis-à-vis* their objects, which means that they can all contribute to a better understanding of the activities of language users.

5. References

Blom, Jan-Petter & John J. Gumperz
1972 Social meaning in linguistic structures: Code switching in northern Norway. In John J. Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics*, 407–434. New York: Holt, Rinehart, and Winston.

Bourdieu, Pierre
1991 *Language and symbolic power*. Cambridge: Polity Press.

Bretxa, Vanessa, Llorenç Comajoan & F. Xavier Vila i Moreno
2016 Is science really English monoglot? Language practices at a university research park in Barcelona. *Language Problems and Language Planning* 40(1). 47–68.

CILAR (Committee on Irish Language Attitude Research)
1975 *Report*. Dublin: The Stationery Office.

Findling, Joav
1971 Bilingual need affiliation and future orientation in extra-group and intra-group domains. In Joshua A. Fishman, Robert L. Cooper & Roxanna Ma (eds.), *Bilingualism in the Barrio*, 337–344. Bloomington, IN: Indiana University.

Fishman, Joshua A.
1964 Language maintenance and language shift as a field of inquiry: A definition of the field and suggestions for its further development. *Linguistics* 2(9). 32–70.

Fishman, Joshua A.
1965 Who speaks what language to whom and when? *La Linguistique* 1(2). 67–88.

Fishman, Joshua A.
1972a The relationship between micro-and macro-sociolinguistics in the study of who speaks what language to whom and when. In John B. Pride & Janet Holmes (eds.), *Sociolinguistics*, 15–32. Harmondsworth: Penguin.

Fishman, Joshua A.
1972b The sociology of language. In Anwar S. Dil (ed.), *Language in sociocultural change*, 1–15. Stanford, CA: Stanford University Press.

Fishman, Joshua A.
1986 Domains and the relationship between micro- and macrosociolinguistics. In John J. Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics: The ethnography of speaking*, 407–434. New York: Holt, Rinehart and Winston.

García, Ofelia & Li Wei
2014 *Translanguaging: Language, bilingualism and education*. Basingstoke: Palgrave Macmillan.

Greenfield, Lawrence & Joshua A. Fishman
1971 Situational measures of normative language views of person, place and topic among Puerto Rican bilinguals. In Joshua A. Fishman, Robert L. Cooper & Roxanna Ma (eds.), *Bilingualism in the Barrio*, 233–253. Bloomington, IN: Indiana University.

Haberland, Hartmut
2005 Domains and domain loss. In Bent Preisler, Anne H. Fabricius, Hartmut Haberland, Susanne Kjaerbeck & Karen Risager (eds.), *The consequences of mobility: Linguistic and sociocultural contact zones*, 227–237. Roskilde: Roskilde University.

Hamel, Rainer Enrique
2008 Les langues des sciences et de l'enseignement supérieur: État actuel et perspectives d'avenir. In Agence Universitaire de la Francophonie (AUF) & Organisation Internationale de la Francophonie (OIF) (eds.), *Séminaire international sur la méthodologie d'observation de la langue française dans le monde*, 193–203. Paris: Éditions OIF.

Harder, Peter (ed.)
2009 *English in Denmark: Language policy, internationalization and university teaching*. Copenhagen: Museum Tusculanum Press.

Hultgren, Anna Kristina
2016 Domain loss: the rise and demise of a concept. In Andrew Linn (ed.), *Investigating English in Europe: Contexts and agendas*, 153–158. Berlin: Mouton de Gruyter.
Hultgren, Anna Kristina, Frans Gregersen & Jacob Thøgersen (eds.)
2014 *English in Nordic universities: Ideologies and practices*. Amsterdam: Benjamins.
Hutton, Christopher
1999 *Linguistics and the Third Reich: Mother-tongue fascism, race and the science of language*. London: Routledge.
Jarvad, Pia
2001 *Det danske sprogs status i 1990'erne med særligt henblik på domænetab* [Status of the Danish language in the 1990s with special reference to domain loss]. Copenhagen: Dansk Sprognævn.
Jørgensen, J. Normann, Martha S. Karrebæk, Lian M. Madsen & Janus S. Møller
2011 Polylanguaging in superdiversity. *Diversities* 13(2). 23–37.
Koufogiorgou, Andromahi
2008 When a dying language becomes a lingua franca. *Multilingua* 27(3). 231–253.
Kristiansen, Tore & Stefan Grondelaers (eds.)
2013 *Language (de)standardisation in late modern Europe: Experimental studies*. Oslo: Novus.
Mackey, William F.
2000 The description of bilingualism. In Li Wei (ed.), *The bilingualism reader*, 26–54. London: Routledge.
Madsen, Mia
2008 *"Der vil altid være brug for dansk": En undersøgelse af 11 naturvidenskabelige forskeres grunde til at vælge henholdsvis dansk og engelsk i deres arbejde* ["There will always be a need for Danish": An investigation of 11 science researchers' reasons for choosing Danish and English, respectively, in their work]. University of Copenhagen: Faculty of Humanities.
Milroy, Lesley
1980 *Language and social networks*. Oxford: Blackwell.
Mortensen, Janus
2017 Transient multilingual communities as a field of investigation: Challenges and opportunities. *Journal of Linguistic Anthropology* 27(3). 271–288.
Oberreuter, Heinrich, Wilhelm Krull, Hans Joachim Meyer & Konrad Ehlich (eds.)
2012 *Deutsch in der Wissenschaft: Ein politischer und wissenschaftlicher Diskurs*. München: Olzog.
Ó Riagáin, Pádraig
1997 *Language policy and social reproduction: Ireland 1893–1993*. Oxford: Clarendon Press.
Parkin, David
1974 Language switching in Nairobi. In William H. Whitely (ed.), *Language in Kenya*, 189–216. Nairobi: Oxford University Press.
Pedersen, Karen Margrethe
2000 *Dansk sprog i Sydslesvig* [The Danish language in Southern Schleswig], 2 vols. Åbenrå: Institut for Grænseregionsforskning.
Pennycook, Alastair & Emi Otsuji
2015 *Metrolingualism*. Abingdon: Routledge.
Preisler, Bent
2005 Deconstructing the 'domain of science' as a sociolinguistic entity in EFL societies: The relationship of English and Danish in higher education and research. In Bent Preisler, Anne H. Fabricius, Hartmut Haberland, Susanne Kjaerbeck & Karen Risager (eds.), *The Consequences of mobility: Linguistic and sociocultural contact zones*, 238–248. Roskilde: Roskilde University.

Rindler Schjerve, Rosita
1996 Domänenuntersuchungen. In Hans Goebl, Peter H. Nelde, Wolfgang Wölck & Zdenek Stary (eds.), *Kontaktlinguistik: Ein internationales Handbuch zeitgenössischer Forschung*, vol. 1, 796–804. Berlin & New York: de Gruyter.
Salö, Linus
2017 *The sociolinguistics of academic publishing: Language and the practices of* homo academicus. New York: Palgrave Macmillan.
Saville-Troike, Muriel
2003 *The ethnography of communication*. Malden, MA: Blackwell.
Schmidt-Rohr, Georg
1933 *Mutter Sprache: Vom Amt der Sprache bei der Volkwerdung*. Jena: Diederichs.
Simonsen, Dag Finn
2002 Å velge bort norsk: Om begrepene »domene« og »domenetap« anvendt på skandinaviske land [Deselecting Norwegian: On the concepts »domain« og »domain loss« applied to Scandinavian countries]. *Norsklæraren* 2002(2). 5–17.
Simonsen, Dag F.
2005 Over the fence – and into English? Reflections on adolescents, academics, linguistic development and language policy in Norway in the early 2000s. In Bent Preisler, Anne H. Fabricius, Hartmut Haberland, Susanne Kjaerbeck & Karen Risager (eds.), *The consequences of mobility: Linguistic and sociocultural contact zones*, 249–271. Roskilde: Roskilde University.
Söderlundh, Hedda
2014 Zooming in on language practices in Swedish higher education: A discussion of five studies and their normative versus dynamic approach towards policy and practice. In Anna Kristina Hultgren, Frans Gregersen & Jacob Thøgersen (eds.), *English in Nordic universities: Ideologies and practices*, 111–126. Amsterdam: Benjamins.
Stickel, Gerhard
2010 Domain loss of a language and its short- and long-term consequences. In Marjeta Huma & Mojca Žagar Karer (eds.), *National languages in higher education*, 13–22. Ljubljana: Založba ZRC.
Torras, Maria Carme & Joseph Gafaranga
2002 Social identities and language alternation in non-formal institutional bilingual talk: Trilingual service encounters in Barcelona. *Language in Society* 31(4). 527–548.
Trudgill, Peter
2003 *A glossary of sociolinguistics*. Oxford: Oxford University Press.
Tsitsipis, Loukas D.
1991 Terminal-fluent speaker interaction and the contextualization of deviant speech. *Journal of Pragmatics* 15(2). 153–173.
Werlen, Iwar
2004 Domäne/Domain. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds.), *Soziolinguistik/Sociolinguistics*, vol. 1, 335–341. Berlin & New York: de Gruyter.

Hartmut Haberland, Roskilde (Denmark)

34. Language contact and education: from language separation to translanguaging pedagogy

1. Introduction: The scope of language contact and education
2. Defining and categorizing bi/multilingual education
3. Types of bi/multilingual education
4. Classroom practices and pedagogical approaches
5. Conclusions
6. References

1. Introduction: the scope of language contact and education

According to Cummins (2011: 5), educational policy makers in the past emphasized the so-called 'effectiveness paradigm', with a view toward "ensuring that students meet universal, one-size-fits-all standards which are assessed by standardized or state-developed tests, all in the service of greater economic competitiveness". Teaching and learning 'majority' languages associated with particular nation-states were key to this homogenizing tendency, thereby imposing 'dominant languages' on children from linguistic minorities. Such "[m]onoglossic language ideologies […] promote the hegemony of monolingualism in a standardized national language [and] suggest that the prestige variety is an objective linguistic category" (Flores and Rosa 2015: 161).

As language contact in teaching settings goes, these ideologies equally fostered a monolingual, separatist approach to language learning that is still prevalent in many classroom contexts. Macaro (2006) connects this to political and economic forces, particularly in relation to the interests of Anglophone countries in spreading the idea that endorsing multilingualism in the classroom is bad practice. Heller and Martin-Jones (2001: 10) argue that the separation of linguistic resources in schools corresponds to the adoption of "a model of language as a property of a nation, that is, a whole and bounded entity that corresponds to a whole, homogenous, and bounded people" which, in turn, leads to the perpetuation of "ideologies that link languages to nations and therefore reproduce ideologically motivated social categories".

While these ideologies may still persist in many policy contexts up until the present day, an ever-growing number of researchers, (language) educators, and policy makers gradually abandoned this monoglossic approach, recognizing, accepting, and embracing the worldwide dominance of societal bi/multilingualism instead (Garcia 2009a; Grosjean 2010). The promotion of various forms of bi/multilingual education, such as immersion programs in Canada in the 1960s and Content Based Instruction in the US in the 1980s, accompanied this new paradigm.

Bi/multilingual education programs worldwide "vary in their goals, language use, and students served, and are also shaped by sociocultural and sociopolitical factors, historical context, as well as the status and power of speakers and languages" (García, Flores, and Homonoff Woodley 2015: 1). As it is beyond the scope of this chapter and handbook to discuss the many different definitions and taxonomies concerned (see Wright, Boun, and García [2015] for a more detailed discussion), I first highlight the scope and character of some of the most influential models and programs in bi/multilingual education which

https://doi.org/10.1515/9783110435351-034

have been developed in recent years, followed by a discussion of the actual classroom practices and pedagogical approaches involved, in particular highlighting language contact phenomena such as classroom code-switching, dynamic bilingualism, and translanguaging.

2. Defining and categorizing bi/multilingual education

Bilingual education has been traditionally distinguished from other types of educational provisions involving two or more languages as "education that aims to promote bilingual (or multilingual) competence by using both (or all) languages as media of instruction for significant portions of the academic curriculum" (Genesee 2004: 548). Multilingual education refers to "the use of two or more languages in education provided that schools aim at multilingualism and multilingual literacy" (Cenoz and Gorter 2015: 2). While both definitions refer to the use and knowledge of two or more languages (and effectively become interchangeable in practice), the first also focusses explicitly on the relationship between language learning and other areas of the curriculum, beyond the process of multilingual language and literacy development *per se*.

Research on bi/multilingual education – "a simple label for a complex phenomenon" (Baker 2006: 213) – is of a highly interdisciplinary nature. Cenoz and Gorter (2015: 6–7) chart the new trends in this field along the lines of "becoming multilingual" and "being multilingual". The first refers to "the process of acquiring communicative competence in a second or additional language", while the latter focuses on "the way multilingual speakers navigate between languages [in] multilingual practices or translanguaging" (Cenoz and Gorter 2015: 6–7).

A well-known classification of bilingual education programs by Baker (2006) refers to the basic distinction between 'strong' and 'weak' forms, depending on whether or not these aim to actively promote bilingualism and biliteracy. Some examples of strong forms would be immersion programs, originally created in Canada, and two-way or dual language programs, characteristically associated with the US In the latter, the idea is that approximately equal numbers of learners from language 'minority' backgrounds and language 'majority' students are enrolled in the same program, and that both languages are used in teaching and learning at different times.

Another important differentiation is that between bilingual education programs targeting 'majority' language speakers and those targeting 'minority' language speakers. These are often aligned with a distinction between models that have an additive versus subtractive intention and outcome. Programs promoting additive bilingualism have been classified as maintenance, heritage, and enrichment programs; all of these aim to extend the speakers' repertoire through adding another language (Hornberger 1991; May 2008). Subtractive models, on the other hand, are designed to replace one language (usually the students' home languages) with another (the dominant society language), reducing the users' linguistic resources; this is typically associated with so-called transitional models or submersion programs.

3. Types of bi/multilingual education

The different possible relationships between languages and content areas in bi/multilingual education, and in particular between language development and academic achievement, defined some of the most frequent (and related) programs in the field, three of which are: (1) language immersion, (2) content-based instruction (CBI), and (3) content and language integrated learning (CLIL). Cenoz (2015: 10, 12) makes crystal clear that these terms overlap and/or are near-synonyms: immersion programs are "the best-known example of CBI" whereas "CLIL and CBI are often considered as two labels for the same reality".

She further points out that "there are no differences between CBI and CLIL regarding their essential properties", and attributes the "accidental differences […] to the specific educational contexts where the programs take place" (Cenoz 2015: 21–22). For the purpose of this handbook, however, it is worth retracing the specific characteristics of each of the three programs.

3.1. Language immersion programs

Immersion programs were developed in Canada from the 1960s onwards, and serve as the classic example of a pioneering initiative to integrate language teaching and learning through the study of academic content. As a result of 'The Quiet Revolution', which led to the prioritization of French over English in the province of Quebec in the 1960s and 1970s (Levine 1991), the province was "[f]aced with the evolving importance of French […] and with evident weaknesses in the current methods of teaching French" (Genesee 2015: 47–48). As a consequence, some English-speaking parents in the community of St. Lambert, near Montreal, began to meet and discuss possible alternative approaches. Despite initial opposition on the part of educational authorities, these meetings eventually gave rise to the establishment of the first French immersion program in 1965, supported by parents and university-based researchers.

Two types of provisions were established. The first took into account the age at which students enter the program: early immersion (4–5 years old), middle or delayed immersion (9–10 years old), and late immersion (secondary or high school level). The second indicated the amount of time spent studying in the immersion (second or foreign) language: total immersion (100 % immersion initially in the second language, leading to 50 % contact at the end of primary school), and partial immersion (50 % immersion in the second language throughout the program).

A separation, or 'sheltered', approach was incorporated into the programs from the beginning. The principle behind this is that learners of the target language are kept apart from native speakers, "at least until their linguistic skills are sufficient to permit them to learn academic content on a par with native speakers" (Swain 1982: 84). In addition, Swain (1983) argued that all children in programs of this type should begin with the same zero level of target language skills. She recommended adopting three basic principles to achieve successful bilingual education: the principle of 'First Things First', or ensuring a sound basis in the child's L1 for L2 learning; the principle of 'Bilingualism through Monolingualism', involving a separation (rather than a concurrent) approach to classroom language use; and the principle of 'Bilingualism as a Bonus', focused on communi-

cating to students the advantages of bilingualism for them, which would lead to a self-fulfilling prophecy.

Since 1965, there has been a steady spread of French immersion programs throughout Canada and beyond. There are now immersion programs in countries as diverse as Australia, Ireland, Paraguay, and Finland. The extensive literature on immersion education research in the 1970s and 1980s in the Canadian immersion settings focused on the development of the learners' L2 proficiency, not least to reassure Anglophone parents that their children would not be disadvantaged. There was no evidence of harmful effects on the development of either the L1 or academic achievement for students in immersion programs: immersion students typically reached the same (or even a slightly higher) proficiency in their L1 as compared to their peers in traditional language teaching modules, in addition to having a far superior mastery of the L2. The overall study results (i.e. for courses other than the language classes) in immersion programs were assessed as slightly higher than those in the traditional education system, thus triggering research about the possible cognitive advantages of bilingual education (see Heller [1990] for an extensive review).

3.2. Content-based instruction

Content-based instruction (CBI) refers to the simultaneous teaching of content and of a target language that acts as a vehicle to learn the content (Lyster 2007). The term is predominantly used in North American contexts and research, and includes a number of different teaching models which have certain similarities to immersion programs. These include Sheltered Content Instruction, a modality where content courses are taught in the target language to a segregated, or 'sheltered', group of learners by a native-speaking content specialist; Adjunct Language Instruction, where students are enrolled in two linked courses, one being a content course and the other a language course; and Theme-based Language Instruction, where the teaching and learning is focused around topics or themes, and which takes place in the language class (Snow 2001).

Much of the scholarly debate among researchers about CBI revolves around the relative importance of language versus content in this approach. Genesee (2004) argues that language is secondary to learning achievement, whereas Met (1998) suggests that students' language proficiency development should be equally as important as their mastery of content.

3.3. Content and language integrated learning

Content and language integrated learning (CLIL) is a more recent approach to content-based language teaching. It was developed and implemented in various European countries, such as Spain, France, The Netherlands, and Belgium (Lasagabaster and Sierra 2010), and is promoted at the level of the European Commission and the Council of Europe. CLIL is defined by Coyle, Hood, and Marsh (2010: 1) as:

> a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language. That is, in the teaching and learning process

> there is a focus not only on content, and not only on language. Each is interwoven, even if the emphasis is greater on one or the other at a given time.

Whereas this definition confirms CLIL and CBI to be near-identical twins, one particular contextual feature which differentiates the schemes is the CLIL-focus on 'culture' as part of the 4Cs framework (Coyle 1999). This framework integrates four contextualized basic concepts defined as: content (subject matter), communication (language learning and use), cognition (learning and thinking processes), and culture (intercultural understanding and global citizenship).

The languages used in European CLIL settings generally include, first, the majority language of the country, or of the region within a country in which the students live, which is also usually their L1. The L2 or target language is almost always English or another major European language, such as French, German, or a language of a majority or minority community of the country in which students live.

CLIL has been widely incorporated into the European context since its origins in the mid-1990s. There have been a number of studies conducted, mainly in Spain, which have shown an emerging interest in the implementation of CLIL practices, particularly in cities where there is more than one official language, such as Basque and Spanish. Furthermore, Lasagabaster (2011) highlights the benefits that CLIL implementations have had on learners' motivation and language competence.

4. Classroom practices and pedagogical approaches

One of the accepted tenets of language use in the earlier stages of CBI (and in immersion programs in particular) was the perceived need for a policy of language separation in the classroom, or in other words: a policy of "bilingualism through monolingualism" (Swain 1983: 40). This has fundamentally changed in recent decades, as the denial of direct language contact in educational settings was challenged and different languages were increasingly used side-by-side in the same classroom setting. This prompted researchers to analyze how "teachers and learners get things done with two languages in bilingual classrooms" (Martin-Jones 1995: 93). Both the analysis of classroom discourse functions, and the actual linguistic behavior of pupils and teachers, proved a fruitful basis for improved understanding of pedagogical and ideological approaches to language contact in classroom settings.

The ensuing section discusses a number of key themes that accompanied that discussion, moving from the mid-to-late 20th century notion of code-switching, to the more recent concepts of dynamic bilingualism, translanguaging, and promotion of linguistic and cultural awareness in educational processes. We will emphasize that an interest in translanguaging as a practice is accompanied by a focus on translanguaging as a pedagogical approach which helps an individual to 'exploit' the different components of his or her linguistic repertoire in an active way, and as such, can be viewed as complementary to language awareness activities.

Even though "[c]odeswitching is, perhaps, the most common, unremarkable and distinctive feature of bilingual behaviour" (Li Wei and Martin 2009: 117), its use and its value in the life of bilinguals often tends to be ignored, downplayed, or even stigmatized

at times as bad practice in classroom settings. Furthermore, the role of the students' L1 is usually ignored in recent methodologies used to teach foreign or second languages, such as communicative language teaching and task-based learning (Cook 2016). Cummins (2008: 65), drawing on the title of Hugh MacLennan's 1945 novel, uses the image of "two solitudes" to convey the lack of contact between the L1 and L2 in traditional language teaching and learning. As Creese and Blackledge (2010) argue with reference to studies carried out in Malaysia and South Africa, teachers and students thus often feel guilty about moving between languages, because code-switching is generally not approved of at the institutional level.

However, there are signs of change. In a study carried out in Jaffna, Sri Lanka, Canagarajah (2011) observed that, while the English teachers, who were locals, ascribed to an institutional English-only ideology (at least in theory), their actual pedagogical practices showed evidence of widespread code-switching between English and Tamil for particular functions, such as building rapport between students and teachers, promoting in-group solidarity, giving commands, and disciplining students. Canagarajah (2011: 207–208) concluded that "[t]he codeswitching revealed how the teachers and learners took on, affirmed, and negotiated different identities, from one context to another, [constituting] a subtle mode of resistance to the agendas of those who control policy in both political and educational domains".

Canagarajah's research shows that – at least in certain settings, such as that studied by Lin (1999) in Hong Kong – there seems to be a movement towards a scenario proposed by Cummins (2008: 65), wherein he argued that, "when we free ourselves from exclusive reliance on monolingual instructional approaches, […] a wide variety of opportunities arise for teaching bilingual students by means of bilingual instructional strategies that acknowledge the reality of, and strongly promote, cross-language transfer". This in turn reinforces Ruiz' (1984) position that all students' languages should be seen as resources, rather than problems, for the development of bilingual or multilingual proficiency in foreign and second language classrooms, and that teachers and administrators should recognize language contact phenomena as a legitimate aspect of language teaching and learning.

There is still, however, a long way to go before such a view will be commonly accepted. In their assessment of the current state of the art in researching classroom code-switching, Lin and Li (2012: 477) identified the "implicit 'legitimating' concerns of researchers [carrying out] studies in the interactional literature" as preventing advances in "analysing how CS practices can be better understood to achieve better pedagogical and social critical purposes".

In 2012, however, there was evidence of a pedagogical and socio-critical emphasis in a study on classroom code-switching carried out by Copp Jinkerson (2012) on an English medium program in a mainstream Finnish language public school in Jyväskylä. She examined the way in which an institutional monolingual norm is produced, oriented to, and resisted in an English-medium first and second grade class located in a mainstream Finnish primary school. Her study, reminiscent in some ways of Canagarajah's (2011), focused on peer interactions in daily classroom life, and was centered on social and cultural interests and links between language socialization, choice, and use in specific local ways, foregrounding the social implications of managing a monolingual norm by examining student agency, resistance, achievement of voice, and uptake of a teacher-like discourse. A summary of the study shows that a common practice adopted by some

of the students observed was sub-teaching, in other words, the use of teacher-like discourses and practices among peers (Copp Mökkönen 2013). Although this practice played an important role in socializing students, particularly among female students, to institutionally preferred monolingual classroom language use and was related to language norms associated with classroom membership, the results also showed evidence of resistance, for example, through the use of mockery by a male student (via intonation and gesture) of another female student's monitoring of classroom language use in English. The practice of sub-teaching could thus also be seen as "a social action accomplished through alignments and positionings [offering possibilities of being] agentively taken-up, co-constructed or resisted" (Copp Mökkönen 2013: 284). Copp Mökkönen (2013: 290) concludes that, "the English monolingual norm plays a complex role in social relationships within the class where the norm may be contested, reformulated or drawn upon in contradictory ways".

An interesting development which has not been mentioned up to this point is the expansion of the notion of language and cultural awareness, particularly in relation to 'minority' language groups. Language awareness is often used "as a synonym for knowledge about language. It broadly refers to an individual's awareness of how language functions as a system and of how language as a social construct is used in society" (Darquennes 2017: 298). While not focused directly on language learning, language awareness activities are meant to encourage pupils to reflect on language as a system as well as on the role of, the attitudes towards and the beliefs about different languages and language varieties in society. Following de Pietro and Matthey (2001: 34–35) and Cummins (2012: 48), and drawing on the French notion of *Éveil aux langues*, 'language awareness' (or in its short form, Evlang), language awareness activities are to be seen as "pedagogical activities centered around languages as 'objects of study' that are meant to trigger the pupils' implicit knowledge about and to foster their susceptibility to differences and similarities between structural features of spoken and written language" (Darquennes 2017: 301). Pupils are brought into contact with different sounds and writing systems that help them to become aware of the arbitrary relationship between sound/form and meaning. Pupils are also encouraged to reflect on the ways in which language diversity manifests itself in society and the classroom, where they have the opportunity to valorize their home languages as well as other languages and language varieties in their individual repertoires. This helps them to affirm their identity as well as the cultural capital of the community they belong to. Evlang is thus "meant to add to the purely linguistic and cognitive skills of pupils (it urges them to reflect on how language works as a system), their sociolinguistic skills (it invites them to reflect on the role of, the attitudes toward, and the beliefs about different languages and language varieties in society), and their psycholinguistic skills (in the sense that pupils are invited to reflect on how the (abstract) knowledge about their own language can help them to familiarize themselves with and/or to acquire other languages)" (Darquennes 2017: 301).

In France, Hélot and Young (2006) documented a language awareness project in an Alsatian primary school in Didenheim, which has had far-reaching effects on the teachers, parents, and children involved in the experience. The researchers also maintain that it "fills a gap in teacher education" (Hélot and Young 2006: 87), in that it helps teachers to recognize that multilingual practices outside school can constitute legitimate classroom resources. They argue that, because language awareness focuses on attitudes rather than on aptitude (a concern of foreign language teaching and learning), this should help

teachers to reflect on the relationships between language and power in the classroom and in the curriculum.

While approaches like Evlang do contribute to fostering positive attitudes toward multilingualism and to accepting language diversity in general, increasing attention is being given to the development of approaches that actually encourage pupils to more actively capitalize on the ways in which bilinguals make use of the different components of their linguistic repertoires. This is where notions such as 'dynamic bilingualism' and 'translanguaging' come into play.

The term 'dynamic bilingualism' is perhaps less well known than 'translanguaging' but in fact, both are closely related. García (2009b: 144) characterizes dynamic bilingualism as "the varying degrees of abilities and uses of multiple language practices needed for people to cross physical or virtual borders". On the other hand, García and Li Wei (2014) maintain that translanguaging is based on the notion that bilinguals have one linguistic repertoire from which they select features strategically to communicate effectively. In other words, translanguaging considers the language practice of bilingual people as the norm. It differs from code-switching, in that code-switching assumes that the two languages of bilinguals are two separate monolingual codes that can be used without reference to each other.

I give examples of some recent studies carried out which provide evidence of how teachers in different contexts have been able to create opportunities that enable students to use all their linguistic resources to access content, negotiate meaning, and express their understanding.

Flores and Schiessel (2014) studied this phenomenon in an educational setting through an ethnographic case study conducted in Grades 3 and 5 in two urban schools in the northeast of the United States. After analyzing data from classroom observations and interviews with the participating teachers, the researchers came to the conclusion that the teachers "used heteroglossic language ideologies to create heteroglossic implementational spaces, though [they were] constrained by the low status given to Spanish instruction produced by a monoglossic educational context" (Flores and Schiessel 2014: 465).

Based on a discussion of multisite ethnographies carried out in two New York schools with an official English as a Second Language designation, García, Flores, and Homonoff Woodley (2015: 221) refer to "the negotiations that take place in schools that construct in-between spaces of promise for immigrant students [building] flexible translanguaging spaces capable of constructing and adjusting to new US bilingual subjectivities". Translanguaging can be thus seen as providing bilingual and multilingual students with spaces which allow for resistance to asymmetrical power balances, characteristic of a monoglossic, separatist school language ideology.

This notion of providing translanguaging spaces is also foregrounded by Martín-Beltrán (2014), who carried out research on how students learning English and students learning Spanish in the US activated multilingual repertoires as they participated in a high school program aimed at promoting reciprocal learning and teaching of multilingual literacy practices. The goal was to establish "how students drew upon Spanish, English, and translanguaging as cultural and cognitive tools to mediate learning in a Third Space" (Martín-Beltrán 2014: 208). The results revealed "fluid and reciprocal affordances for language learning during interactions among linguistically diverse peers as they draw upon translanguaging practices" (Martín-Beltrán 2014: 208).

Increasingly, researchers are becoming interested in connecting this phenomenon in teaching and learning with ways of assessing translanguaging practices in the classroom. Flores and Schiessel (2014: 472), for example, extended their discussion to questions of the reform of common core state standards, arguing that "[i]t is only when assessments are designed in ways that affirm the dynamic language practices of emergent bilingual students that teachers can truly begin to provide spaces that allow and even encourage students to use their entire linguistic repertoire in ways that empower them while providing them with the skills they need to succeed in the 21st century".

In this respect, Otheguy, García, and Reid (2015: 299) argue: "Accuracy of measurement is a bedrock value in the context of educational testing. Yet forbidding bilinguals to translanguage, or assessing it negatively, produces an inaccurate measure of their language proficiency". For their part, Menken and Shohamy (2015: 421) maintain that "the use of multilingual assessments significantly contributes to higher scores on academic tasks and more accurately reflects the knowledge of multilingual test takers". López, Turkan, and Guzman-Orth (2014) claim that flexible mathematics bilingual assessment − that allows students to translanguage whenever needed − has a positive impact on the students' skills, their abilities in mathematics, and their bilingual language use. Ossa Parra (2018) conducted ethnographic research in collaboration with a third grade Sheltered English Instruction teacher in a Boston public school serving a predominantly Latino population. The results indicated how translanguaging pedagogy can be planned and enacted in ESL instruction with third grade bilingual students with a varied range of English and Spanish proficiencies. The flexible language practices promoted during these lessons supported students' meaning making about texts and their salient language features in the areas of semantics, morphology, and syntax, as well as affirming their identities as competent readers and language users.

5. Conclusions

Over the last 50 or so years, there have been dramatic developments in the conceptualization and study of language contact. The field of education is no exception to this rule. From a purist, monolingual vision, which entailed the strict separation, at least in theory, between the languages taught in the curriculum and the languages of the learners, there has been increasing recognition of the value of both teachers' and learners' bilingual and multilingual repertoires in teaching and learning processes. This recognition has involved a reevaluation of hybrid language use in the classroom. It has also entailed a growing understanding of the complexities involved in becoming bi/multilingual speakers and writers in educational settings. Sierens and Van Avermaet (2014: 220) observe that, "when schools acknowledge and use the multilingual repertoires of the students, they present them with better chances of educational success".

However, as we have seen, there is often evidence of tension generated between monolingual institutional language policies and bi/multilingual classroom practices, which in turn affects teachers in their classrooms. As Cummins (2011: 8) notes, "educators themselves frequently have internalized the discourse of effectiveness and are unprepared for the realities of highly diverse classrooms. They see problems rather than opportunities in their students' linguistic and cultural differences". There is a case, then, to be

made for studies to be carried out on the pedagogical and social implications of bilingual and multilingual use in the classroom involving language contact. This would, as Cummins (2011: 5) recognizes, allow us to acknowledge and embrace "the reality of diversity, not as a problem to be managed or resolved, but rather as an opportunity to enrich the lives of students and teachers".

6. References

Baker, Colin
2006 *Foundations of bilingual education and bilingualism*, 4th edn. Clevedon: Multilingual Matters.

Canagarajah, Suresh
2011 Translanguaging in the classroom: Emerging issues for research and pedagogy. *Applied Linguistics Review* 2. 1–28.

Cenoz, Jasone
2015 Content-based instruction and content and language integrated learning: The same or different? *Language, Culture and Curriculum* 28(1). 8–24.

Cenoz, Jasone & Durk Gorter
2015 Towards a holistic approach in the study of multilingual education. In Jasone Cenoz & Durk Gorter (eds.), *Multilingual education: Between language learning and translanguaging*, 1–15. Cambridge: Cambridge University Press.

Cook, Vivian
2016 *Second language learning and language teaching*, 5th edn. Routledge.

Copp Jinkerson, Alicia
2012 *Socialization, language choice and belonging: Language norms in a first and second grade English medium class* (Jyväskylä Studies in Humanities 182). Jyväskylä: University of Jyväskylä dissertation.

Copp Mökkönen, Alicia
2013 An ethnographic study of language socialization and choice in a first and second grade English medium classroom in Finland. *Journal of Immersion and Content-Based Language Education* 1(2). 279–295.

Coyle, Do
1999 Theory and planning for effective classrooms: Supporting students in content and language integrated learning contexts. In John Masih (ed.), *Learning through a foreign language*, 46–62. London: CILT.

Coyle, Do, Philip Hood & David Marsh
2010 *CLIL: Content and language integrated learning.* Cambridge: Cambridge University Press.

Creese, Angela & Adrian Blackledge
2010 Translanguaging in the bilingual classroom: A pedagogy for learning and teaching. *Modern Language Journal* 94(1). 103–115.

Cummins, Jim
2008 Teaching for transfer: Challenging the two solitudes assumption in bilingual education. In Jim Cummins & Nancy Hornberger (eds.), *Encyclopedia of language and education*, 2nd edn., vol. 5: Bilingual education, 65–75. New York: Springer.

Cummins, Jim
2011 Preface. In Anne-Marie de Mejía & Christine Hélot (eds.), *Empowering teachers across cultures: Enfoques críticos, perspectives croisées*, 5–8. Frankfurt am Main: Peter Lang.

Cummins, Jim
2012 Language awareness and academic achievement among migrant students. In Claudine Balsiger, Dominique Bétriz Köhler, Jean-François de Pietro, Christiane Perregaux (eds.), *Éveil aux langues et approches plurielles*, 41−54. Paris: L'Harmattan.
Darquennes, Jeroen
2017 Language awareness and minority languages. In Jasone Cenoz, Durk Gorter & Stephen May (eds.), *Encyclopedia of language and education: Language awareness and multilingualism*, 3rd edn., 297−308. New York: Springer.
de Pietro, Jean- François & Marinette Matthey
2001 L'éveil aux langues: des outils pour travailler la différence. *Langage & Pratiques* 28. 31−44.
Flores, Nelson & Jamie L. Schiessel
2014 Dynamic bilingualism as the norm: Envisioning a heteroglossic approach to standards-based reform. *TESOL Quarterly* 48(3). 454−479.
Flores, Nelson & Jonathan Rosa
2015 Undoing appropriateness: Raciolinguistic ideologies and language diversity in education. *Harvard Educational Review* 85(2). 149−171.
García, Ofelia
2009a *Bilingual education in the 21st century: A global perspective.* Oxford: Wiley-Blackwell.
García, Ofelia
2009b Education, multilingualism and translanguaging in the 21st century. In Tove Skutnabb-Kangas, Robert Phillipson, Agit K. Mohanty & Minati Panda (eds.), *Social justice through multilingual education*, 140−158. Bristol: Multilingual Matters.
Garcia, Ofelia & Li Wei
2014 *Translanguaging: Language, bilingualism and education.* New York, NY: Palgrave MacMillan.
Garcia, Ofelia, Nelson Flores & Heather Homonoff Woodley
2015 Constructing in-between spaces to 'do' bilingualism: A tale of two high schools in one city. In Jasone Cenoz & Durk Gorter (eds.), *Multilingual education: Between language learning and translanguaging*, 199−224. Cambridge: Cambridge University Press.
Genesee, Fred
2004 What do we know about bilingual education for majority language students? In Tej K. Bhatia & William Richie (eds.), *Handbook of bilingualism and multiculturalism*, 547−576. Malden: Blackwell.
Genesee, Fred
2015 Canada: Factors that shaped the creation and development of immersion education. In Peeter Mehisto & Fred Genesee (eds.), *Building bilingual education systems: Forces, mechanisms and counterweights*, 43−58. Cambridge: Cambridge University Press.
Grosjean, François
2010 *Bilingual life and reality.* Cambridge, MA: Harvard University Press.
Heller, Monica
1990 French immersion in Canada: A model for Switzerland? *Multilingua: Journal of Cross-Cultural and Interlanguage Communication* 9(1). 67−85.
Heller, Monica & Marilyn Martin-Jones
2001 Introduction: Symbolic domination, education and linguistic difference. In Monica Heller & Marilyn Martin-Jones (eds.), *Voices of authority: Education and linguistic difference*, 1−28. Westport, Connecticut: Ablex Publishing.
Hélot, Christine & Andrea Young
2006 Imagining multilingual education in France. In Ofelia García, Tove Skutnabb-Kangas & Maria Torres-Guzmán (eds.), *Imagining multilingual schools*, 69−90. Clevedon: Multilingual Matters.

Hornberger, Nancy
1991 Extending enrichment bilingual education: Revisiting typologies and redirecting policy. In Ofelia García (ed.), *Bilingual education: Focusschrift in honor of Joshua A. Fishman*, 215–234. Amsterdam: John Benjamins.
Lasagabaster, David
2011 English achievement and student motivation in CLIL and EFL settings. *Innovation in Language Learning and Teaching* 5(1). 3–18.
Lasagabaster, David & Juan Manuel Sierra
2010 Immersion and CLIL in English: More differences than similarities. *ELT Journal* 64(4). 367–375.
Levine, Mark
1991 *The reconquest of Montreal*. Philadelphia: Temple University Press.
Li Wei & Peter Martin
2009 Conflicts and tensions in classroom codeswitching. *International Journal of Bilingual Education and Bilingualism* 12(2). 117–122.
Lin, Angel M. Y.
1990 *Teaching in two tongues: Language alternation in foreign language classrooms. Research Report 3.* Hong Kong: City Polytechnic of Hong Kong.
Lin, Angel M. Y.
1999 Doing-English-lessons in the reproduction or transformation of social worlds? *TESOL Quarterly* 33(3). 49–84.
Lin, Angel Y. M. & David C. C. Li
2012 Codeswitching. In Marilyn Martin-Jones, Adrian Blackledge & Angela Creese (eds.), *The Routledge handbook of multilingualism*, 470–481. London: Routledge.
López, Alexis A., Sultan Turkan & Danielle Guzman-Orth
2014 A study on the use of translanguaging to assess the content knowledge of emergent bilingual students. Paper presented at the Invited Colloquium, Negotiating the Complexities of Multilingual Assessment, at the annual American Association for Applied Linguistics (AAAL) Conference, Portland, OR, 18–21 March.
Lyster, Roy
2007 *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam: John Benjamins.
Macaro, Ernesto
2006 Codeswitching in the L2 classroom: A communication and learning strategy. In Enric Llurda (ed.), *Non-native language teachers: Perceptions, challenges and contributions to the profession*, 63–84. New York: Springer.
Martín-Beltrán, Melinda
2014 "What do you want to say?" How adolescents use translanguaging to expand learning opportunities. *International Multilingual Research Journal* 8(3). 208–230.
Martin-Jones, Marilyn
1995 Codeswitching in the classroom: Two decades of research. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, 90–112. Cambridge: Cambridge University Press.
May, Stephen
2008 Bilingual/immersion education: What the research tells us. In Jim Cummins & Nancy Hornberger (eds.), *Encyclopedia of language and education*, 2nd edn., vol. 5: Bilingual education, 19–34. New York: Springer.
Menken, Kate & Elana Shohamy
2015 Invited colloquium on negotiating the complexities of multilingual assessment, AAAL Conference 2014. *Language Teaching* 48(3). 421–425.

Met, Myriam
1998 Curriculum decision-making in content-based language teaching. In Jasone Cenoz & Fred Genesee (eds.), *Beyond bilingualism: Multilingualism and multilingual education*, 35–63. Clevedon, UK: Multilingual Matters.
Ossa Parra, Marcela
2018 *"En español es distinto": Translanguaging for linguistic awareness and meaningful engagement with texts.* Boston: Boston College dissertation.
Otheguy, Ricardo, Ofelia García & Wallis Reid
2015 Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review* 6(3). 281–307.
Ruiz, Richard
1984 Orientations in language planning. *NABE: The Journal for the National Association for Bilingual Education* 8(2). 15–34.
Sierens, Sven & Piet Van Avermaet
2014 Language diversity in education: Evolving from multilingual education to functional multilingual learning. In David Little, Constant Leung & Piet Van Avermaet (eds.), *Managing diversity in education: Languages, policies, pedagogies*, 204–222. Bristol: Multilingual Matters.
Snow, Marguerite Ann
2001 Content-based and immersion models for second and foreign language teaching. In Marianne Celce-Murcia (ed.), *Teaching English as a second or foreign language*, 3rd edn., 303–318. Boston: Heinle & Heinle.
Swain, Merrill
1982 Immersion education: Applicability for non-vernacular teaching to vernacular speakers. In Beverly Hartford, Albert Valdman & Charles Foster (eds.), *Issues in international bilingual education: The role of the vernacular*, 81–97. New York: Plenum Press.
Swain, Merrill
1983 Bilingualism without tears. In Margaret Clarke & Jean Handscombe (eds.), *On TESOL '82: Pacific perspectives on language learning and teaching*, 35–46. Washington, D.C.: Teachers of English to Speakers of Other Languages.
Wright, Wayne E., Sovicheth Boun & Ofelia García (eds.)
2015 *The handbook of bilingual and multilingual education*. Malden, MA: Wiley-Blackwell.

Anne-Marie de Mejía, Bogotá (Colombia)

35. Media and contact linguistics

1. Introduction
2. Three perspectives on media and linguistic contact
3. On media and linguistic contact
4. References

1. Introduction

Media, as extensions of our senses and creators of the environment in which we live (McLuhan 1994), shape, and are in turn shaped by, our own linguistic production. Tradi-

https://doi.org/10.1515/9783110435351-035

tional mass media (i.e. the press) set out to distribute information or entertainment, while simultaneously promoting a unification and standardization of certain linguistic varieties in both spoken and written forms of language. Although mass media outlets were not always developed with the intent of becoming institutions responsible for the standardization of language, they are considered models for correct speech practices and spaces for the use of formal written language. In this respect, they joined other institutions, such as schools, religious groups, and governmental bodies, in granting superior value to a certain way of using any given language.

Although media outlets were developed to reach wider audiences, they tended towards a linguistic simplification which developed specific registers drawn from linguistic continua. In a sense, media − from books onwards − created a separation that was nonexistent in language usage. They broke the continuum and developed specific boundaries for each language in play. Thus, alongside political structures and religious bodies that determined official languages, media created the limits of many of these different varieties and produced specific registers that were granted more social value.

Based on the interest of reaching those audiences, they also privileged the language of those in power, who primarily were also part of the linguistic majority, thus turning most media products into monolingual spaces where only one language was used at a time, despite multilingual production spaces being common. However, in the same way that the 'one-nation, one-language' approach to state-building revealed fractures in the political space (Castells 2005) and was challenged by national minorities and immigration, media as monolingual spaces have also been challenged, and multilingual media outlets exist all over the world. Thus, not only have media worked as creators of language discrimination and standardization but have also become sites of linguistic interaction. Media products may show evidence of linguistic contacts occurring in their broadcast or distribution spaces, thus bringing the linguistic contact to media consumers and audiences. As media in various languages cross boundaries, thanks to commercial and information technology networks, cultures and languages become hybridized (García Canclini 2001). Furthermore, production teams worldwide, as in most transnational and global corporations, are constituted of people with large linguistic repertoires (Angouri 2013), and the process of developing media products includes various instances of language debates and negotiations, which tend to include the transfer of linguistic qualities from one language to the next. Code-switching and lexical borrowings are common in most media products, as much a part of bound novels as of song files, Twitter discussions, and television broadcasts.

Media products in certain languages have also made their way across borders, reaching people who may have the cultural capital to understand them. In cases where they do not, unintelligible, but catchy or attractive, elements from other languages are incorporated into a linguistic group, where those lexical elements or turns of phrase become associated with glamor, good taste, and/or social recognition (Blommaert 2010). At the same time, through the use of the Internet and social media networks, more linguistic elements are exchanged globally in cyberspace, providing another space where languages meet and interact. This led Androutsopoulos (2015) to coin the term *networked multilingualism* to account for some practices in computer-mediated communication. This is relevant because "throughout human history, most language contacts have been face to face (…) [but] there are other possibilities, especially in the modern world with a novel

means of worldwide travel and mass communication: many contacts occur through written language only" (Thomason 2013: 32).

This brief description of the current state of affairs at the intersection between media and languages serves to point out the many instances where languages come into contact, and where that contact becomes relevant for crosspollination and creative linguistic development. In particular, "interest in mass media, especially television, in the context of language change is grounded both in the social function of mass media and in properties of media language" (Androutsopoulos 2014c: 7). Language contact may take place in at least three ways: between media products and a given linguistic community, as in the internationalization of television products or rock songs; within media products themselves, where languages may coexist, and code-switching and/or the exchange of loanwords become evident; or in the spaces and processes of media production, including interactions in international production teams, Internet forums, or Massive Multiplayer Online Games. These three options are interesting avenues of research. First, they address the longstanding debate about media influence on audiences. Language contact through media may contribute to processes of linguistic 'hygiene' (see Cameron 2012), that is, the process of setting boundaries of a normalized and selected register of a given language. Contact may also further the inclusion of linguistic borrowings or neologisms. Second, they identify the media at large as the locus of language contact, representing language contact and interaction. Third, they showcase how media corporations, such as businesses or community endeavors, also become specific sites of language contact, code-switching, and linguistic exchange. In combination with the fact that media are no longer centralized or unidirectional, turning audiences into users, and consumers into producers (Jenkins, Ito, and Boyd 2016), these new sites of communication become new and interesting spaces of linguistic contact by default (Androutsopoulos 2015).

The various processes of interaction between media and language are increasingly being analyzed against the backdrop of sociolinguistic theories of language change (Androutsopoulos 2014b). Stuart-Smith (2017: 29) foregrounds the variationist framework, in particular, as an alternative to interactional linguistics to frame the media and language relationship, given its underlying rationale of "quantify[ing] aspects of language (the linguistic variable) with aspects of society (the sociolinguistic variable)". Given that she highlights that "for language we need to adopt a more nuanced understanding of the notion of media influence" (Stuart-Smith 2017: 32), dedicated effort is required to develop a variety of methodological tools and strategies to understand the interplay of media and linguistics beyond the anecdotal.

In this chapter, emphasis is placed on using examples from minority languages in contact with majority languages, particularly because their situation often showcases most of the topics addressed here. The impact of majority languages upon minority languages is evident during linguistic contact, and code-switching is a common occurrence and an issue of concern, particularly for language purists. However, some other examples are drawn from the study of dialectal varieties and from the influence of a global language, such as English. All these instances provide a broader view on the matter of contact linguistics and the media.

2. Three perspectives on media and linguistic contact

Media linguistics is an area of study concerned with the use of language in media products or processes. It can be narrowed down to a research endeavor that "uses data from media settings to answer research questions raised by linguistics itself, such as language change in everyday contexts" (Perrin 2017: 11). It intersects with media studies in the interest it places on studying media products as well as development and production processes, and by focusing on the specific ways in which media carry out linguistic practices.

We can address the relationship between media outlets − such as books, newspapers, radio shows, films, TV shows, blogs, vlogs, and social media posts − from at least three different perspectives. One prominent perspective, often called 'localization' by those in translation studies, is to examine how that product is inserted into a cultural market which uses a different language (or a specific version of a language), and how that language (or language version) is replicated or incorporated into practice by a community. Research in this area relates to traditional studies of media effects, since it concerns itself with how given traits of the media products influence the language use of audiences or media users. Another perspective is to focus on the product, the given language(s) which are used in it, and their interaction. The final (two-faced) perspective includes studying the site of media production, whether it is a physical place − such as studies of newsroom multilingualism (Perrin 2017) − or a digital one − such as an Internet platform like Facebook (Androutsopoulos 2015). These three perspectives are explored in more detail below.

2.1. Media products, linguistic purity, standardization, and code-switching

Although research on media influence and linguistic development has been unable to trace a clear causal relationship between the two (Schmidt 2015; Selnow and Bettinghaus 2009; Stuart-Smith 2007), the process of *verbal hygiene* (Cameron 2012; Paffey 2007) has often been discussed in contexts where media attempt to establish or construct a specific variety or register of a given language in opposition to the continuum where the language thrives. It is expected of many media that they present forms of the language that are understandable to large populations by avoiding specific dialectal forms, or whenever there is a need to use those forms, to make their use explicit or clear. The standard language used in television broadcasts was often considered to be the most proper form in the use of said language, and since people in the media seem to be culturally relevant, they create a virtuous circle through which that specific register is considered to be the best form of said language (Bell 1983). This is particularly salient in the case of minority language broadcasts. For instance, in the case of the Welsh language TV channel S4C, Jones (2013) highlights how there are linguistic policies in place demanding the language used in Welsh language programming to be clear and correct. Similarly, Mano (2006: 296−98) discovered that listeners to Radio Zimbabwe's broadcast in Shona were very happy with the use of the standardized Harare-Zezuru dialect, considering that version to be the most understandable and clear for the audience.

In the Basque country, a broadcast version of Euskara Batua was developed from the standardized version of written Euskara based on newspaper journalists' language use (Díaz Noci 2002).

Although in all cases presented here there is the goal of creating or using a standardized variety of the language for broadcast, Jones (2013), Mano (2006), and Díaz Noci (2002) document that media outlets use different varieties of minority languages as well (see also Chapter 61). That is to say, no matter the interest in providing a standardized version of the language, they need to accept that some media participants − in radio and television broadcasts − may employ other dialects. This is probably because "the standardization process can lead to a situation where hardly anyone recognizes their language or linguistic practices in the media − the language practices of the media are too far apart from the 'messy' and emergent multilingual practices of the community" (Pietikäinen and Kelly-Holmes 2011: 65). In these cases, normative forces exert pressure for certain linguistic changes, fostering linguistic ideologies while denying space for (or removing) other languages, structures, and forms (Ramallo 2017). This situation may take place by defining the appropriateness of a given variety of a language, and by creating a requirement for the use of a specific orthography, tonality, and register in order to be socially acceptable.

Linguistic contact may take place, then, within one language community involving the different varieties that come into play in traditional media production. For instance, in Nasa territory within Colombia, a radio broadcast in a dialect of the indigenous Nasa Yuwe language that was not easily understandable for other members of the Nasa led to debates about standardized media forms (Uribe-Jongbloed and Peña Sarmiento 2014). In a similar way, there have been ample debates regarding media in Galician which are concerned with the linguistic quality of the language in the media products (Ramallo 2017). As these examples highlight, media products become elements in the contact between two (or more) varieties of a given language, and may lead to the same discussions often found taking place on media in multilingual areas.

In fact, one of the main arguments for the development of minority language media is that they provide a domain for the language in a media sphere controlled by majority language media. The idea here is that majority language media promotes language shift away from the minority language, and media in minority languages would at least mitigate that advance. Much minority language media advocacy is predicated upon this principle, and "while media alone cannot revitalize a minority language, they are important catalysts in changing the ideological and functional value attached to lesser-used minority languages" (Moriarty 2009: 147).

On the other hand, media products developed in languages different from the local (or majority) language where the media products are found may be a form of linguistic contact leading to influence on users' language skills. Radio stations in a foreign language may be sites that promote language contact and learning, as Tsuda and Lafaye (2005) noticed in the case of RADIO-I in Nagoya, Japan. In a globalized setting, where media in global English are the norm, whether through local subtitles or dubbing into a specific variety (for instance, into the Latin American standard version of Spanish), media products become the purveyors of language contact. Returning to the debate mentioned earlier about the impact of direct media exposure on language acquisition, and thus on language shift, there is still no proof of direct impact. That being said, some products which include specific varieties of the language might have some incidental or

short-term impact, as may be the case in the use of certain features of the Cockney dialect by Glasgow youth following exposure to *Eastenders* (Stuart-Smith et al. 2013), in the use of Anglicisms to talk about sports in France (Bernard-Beziade 2008), or in the use of words or elements from a variety of languages in French *bandes dessinées* (Ben-Rafael and Ben-Rafael 2012).

However, in the opposite direction, it has been amply demonstrated that the use and consumption of diaspora media and other linguistic minority media is helpful for maintaining and extending linguistic competence in an L1 under the influence of an L2, such that "for minority groups, greater exposure to and use of media in the L1 has the potential to promote communication with a larger ingroup, which would positively influence group interactions and foster cultural markers" (Gaudet and Clément 2009: 224). This is also the case for those who maintain a heritage language as a diaspora, as in the case of maintaining Mandarin knowledge and competence in the US, where Zhang (2009: 207) found that "though those [heritage language] media might not be deliberately used for maintenance of speech and literacy level, they constitute an ecology that is conducive to such a purpose". These results are similar to those of the Spanish speakers in the US which continue consuming music and engaging in texting in Spanish (Velázquez 2017). Media, as part of the domains surrounding L1 users, may delay attrition and language shift. Television, for instance, could create a geolinguistic region which overcomes borders and allows for an atomized audience to remain in contact with a language (Amezaga et al. 2012).

Media products may themselves be places for, or examples of, linguistic contact, be it through phone interviews in a language different from the usual programming, by presenting characters from various linguistic backgrounds on a TV drama series, or in multilingual debates on Facebook walls. Such interactions provide situations which may lead to the production of new terminology and the use of loanwords. For instance, the use of English mixing in Korean TV commercials promotes a modernity associated with the European language, in such a way that "for modern Koreans, linguistic hybridity − codemixing in English − resolves the tension between global (i.e., dominant English and American culture) and local practices" (Lee 2006: 87). In pop music, access to English mixing also allows for "young Korean-English bilinguals to freely express socially repressed feelings such as sexual desires, derogatory expressions, and anger" (Lee 2006: 87). Concomitantly, due to the presence of English mixing in a variety of media products in Korea, and in other institutions such as businesses and schools, Fayzrakhmanova (2016: 14) concludes that "English is now a Korean language as well, in the sense that English can be served to express a unique Korean culture, in the sense that English is actively used for implementing Korean communicative purposes and in the sense that English is able to express and transmit Korean identity and values".

Media products become sites of research: places where code-switching can be actively studied. It could be code-switching in diaspora radio, as in the case of Gujarati radio broadcasts in the UK (Zipp 2017), use of minority and majority languages by a minority language broadcaster (Atkinson and Kelly-Holmes 2011), or diachronic development of Colloquian Belgian Dutch in comic books (Meesters 2012). Atkinson and Kelly-Holmes (2011) take a direct look at code-switching between Irish and English in a radio comedy, to elucidate how the conscious creation of this code-switching reflects the social perspective of the language and positioning of the Irish language *vis-à-vis* English. They conclude that "traditional attitudes towards use and ownership of Irish are in fact reinforced

in a genre which in Ireland is often portrayed as celebrating and promoting the Irish language", and that the perspective of the sketches is "linguistically reactionary rather than progressive" (Atkinson and Kelly-Holmes 2011: 259). A similar argument can be made with the case of *Mataku*, the TV series that sought to put the Maori worldview into a global audiovisual market (Glynn and Tyson 2007). Although Glynn and Tyson (2007) say very little about the code-switching and multilingual aspects of the show, it becomes immediately apparent that the interest is in changing attitudes through the presence of the minority language. This both recognizes the coexistence of English and *Te Reo* Māori, and presents the possibility of exchange between the two.

In the aforementioned Welsh case of S4C, Jones (2013) also defines the level of permeability − the fact that English terms might appear on Welsh language broadcast − of the channel's linguistic policies, including those pertaining to their websites and social media outlets. She cites many of the guidelines of S4C to show how they are conscious of the inevitability of English code-switching and terminology arising in the media production processes. Regarding the guidelines, Jones (2013: 40) concludes that "[t]hese guidelines demonstrate an understanding of the complexities of producing television programmes and providing a broad television service in a minoritised language on a single channel. They reflect an awareness of how the specific medium of television can be used to communicate with an audience that is actively acknowledged to be heterogenous in term of language skills". This view is in line with the research undertaken with Basque, Irish, and Sámi media, where Kelly-Holmes, Moriarty, and Pietikäinen (2009: 239) found that the policies in practice were more fluid and polycentric, and they "point to a shift from multilingualism, as conceived in official policy as a type of parallel monolingualism, to an understanding of multilingualism as lived heteroglossia". In other words, media recognize the multilingual repertoires of their audiences and accommodate them through their products and output. This was also the case in San Andres Creole radio production, where people were allowed to express their views in whichever language they chose during phone-ins, although Creole was the main language of discussion (Uribe-Jongbloed 2016). The permeability of media outlets to include other languages and allow for code-switching might be viewed by those on the normative end as detrimental to a given language, yet it instead exemplifies the reality of the multilingual settings where languages are being used.

By studying media products diachronically, as Meesters (2012) presents in his study of comic books, one can research developments in linguistic use and the history of the language. This variationist approach seems to be an avenue to comprehend the effects of both linguistic ideologies, as presented by de Bres and Belling (2015), and historical modifications that might stem from standardization or vernacularization practices.

2.2. Multilingual teams of production

Another space where contact linguistics has found an area of study is in the constitution of multilingual groups of media production. In minority language media and in transnational media corporations and projects, the sites of media production are spaces where code-switching between languages take place. Journalists work in several languages when shifting between searching, interviewing, editing, and publishing research. As pre-

sented by Catenaccio et al. (2011), studying the end results of media multilingual production eclipses the processes which inform and determine decision-making strategies. To study media news production in multilingual settings implies examining the entire process, not exclusively its end result.

Similarly, observation of media teams working on Welsh- and English-language audiovisual products revealed elements of code-switching, language preferences for English for research purposes, and extensive use of English-language terminology for the production of live television (Uribe-Jongbloed 2016). For the purposes of studying linguistic decision making and language decisions in journalistic practices, Perrin (2017) has developed a set of matrices that may be used to gain better insight into the process. It serves to disentangle the complex processes that take place in the selection and decision-making activities of multilingual journalists.

Further work on media practices (in contrast to media products) might also help to highlight the different ideologies that may continue to exist regarding languages in contact and their relationships to one another. In the multilingual workplace of global communication and media companies, this is certainly an avenue of interest that could extend beyond intercultural studies, where it usually remains.

2.3. Languages in digital social media

Unlike traditional print and broadcast media, digital social media (including the Internet) enable more direct, quicker interactions between people with a variety of linguistic repertoires, and allow us to study the linguistic effects involved. Although most studies using information from digital social media networks are still developing the most appropriate methodological tools to address debates about language contact, it is undoubtedly a new domain (or space) where language contact processes can be studied.

Zhuravleva, Bot, and Hilton (2016) created a methodology designed to capture information from users of Facebook regarding linguistic use. While their results were largely similar to those obtained from traditional questionnaire data on self-reported language choice and use, they were able to provide new insights that could be further exploited. For their method, termed the social media method (SMM), they had students from diverse L1 backgrounds monitor Facebook messages posted by at least six of their friends in public discussions, and register their language use. The researchers compared the results with a self-report questionnaire from the same participants. They concluded that the SMM was as reliable as the self-administered questionnaire, yet they acknowledge the difficulties in setting up such a study, and the potential privacy concerns it may cause. Similarly, Patro et al. (2017) also aimed at developing a computational method that could analyze and predict borrowed words and instances of code-switching through Twitter. Although their methods require replication, they offer an interesting way of exploring social media to study phenomena related to linguistic contact. Finally, Aragrande (2016) has devised a theoretical-methodological approach combining critical discourse analysis, corpus linguistics, and translation studies to study multilingual audiovisual and written news. All these strategies are worth replicating to strengthen the validity of their methodological construct.

Most studies in this vein limit their scope to specific digital media products or small group samples. For instance, in a study of the revitalization of Mayan, Cru (2017) found that urban youth create rap songs in Mayan, or Mayan mixed with Spanish, and upload them to YouTube and/or Facebook. In the comments section, the Yucatan youth tend to include the lyrics of their songs, which are often misspelled with Spanish homophones. Regarding this phenomenon, Cru (2017: 487) states that "this is unsurprising as informal digital environments such as YouTube do allow for deviation of standard orthographies". Facebook can also be used as a site to study code-switching using, for example, some of the groups created by diaspora families, as in the case of Arabic-English code-switching among migrants to Australia (Alhazmi 2015). De Bres and Belling (2015) concentrated on analyzing the *de facto* linguistic policy that arises in a Facebook group in Luxemburg, presenting some of the linguistic ideologies that transformed a multilingual space into one with Luxemburgish as a clear dominant language. They describe how the Facebook environment enables a diachronic study, where posts can be traced chronologically. Although their approach is limited to a case study, the possibilities for the diachronic study of language contact remain attractive. However, there are privacy concerns which rightly limit the full potential of exploration of Facebook and other digital social networking sites (Zhuravleva, de Bot, and Hilton 2016), in particular when focusing on personal homepages or accounts of individual language users.

If traditional broadcast and print media were the keepers and promoters of standardized versions of a language, digital social media networks seem to provide opportunities to turn to other varieties, spellings, and grammars. They seem to be evidence for the vernacularizations of language in the media, something that Coupland (2014) has argued as well for broadcast media. It becomes important to recognize that "we should also remember that on sites such as Facebook, YouTube, Second Life, as well as the discussion boards of the BBC Voices project, people can play with language in an environment with little regulation or censorship about the content uploaded online, particularly in relation to language choice" (Kelly-Holmes and Milani 2011: 481).

Social media sites provide a new space to study language contact, particularly for those languages where there is ample presence. It is also an area that takes advantage of the availability of linguistic contact, which can be mined through the development of specialist software. Some remaining issues must be borne in mind, including debates about privacy, the age groups which are more likely to participate, and the distribution of debates in relation to the real world. It is, nonetheless, an area of research worth pursuing.

3. On media and linguistic contact

Media are changing as part of the process of media convergence, and those changes also require innovation in research practices related to linguistic contact. The change from a broadcasting structure to a more participative one challenges notions of linguistic standardization, and renders it more difficult to frame traditional media as keepers of the standardized version of the language (Coupland 2014). Any process of standardization "is strongly influenced by competing beliefs and ideas about language in a community" (Darquennes and Vandenbussche 2015: 7), and is created and fostered, while also disput-

ed and discussed openly, through media. Our understanding of media influence on linguistic practices remains incomplete because, like many other forms of media influence, it is almost impossible to disentangle media consumption from other activities of social interaction.

Sociolinguistic approaches are currently taking all of these perspectives into account, following Androutsopoulos' (2014a) recommendation to integrate insights from other disciplines or areas of inquiry as indispensable tools for understanding sociolinguistic change in societies harboring a variety of media exchanges.

The three perspectives presented are intended as 'avenues of research' that can push the discussion about the interplay between media and language forward, against the backdrop of the variationist framework within sociolinguistics. Stuart-Smith (2017: 39) states that "the review of the literature suggests that actual media-induced changes on language are relatively unusual. This point also aligns with what has been discovered and theorized in different branches of mass communications research, namely that the media contribute to rather than impose, and that influence is mainly about what the viewer brings to the media".

This conclusion aligns with the perspective of a 'mediatized' relationship between media and language. Despite the lack of certainty about media impact, there is enough evidence to show that there is some form of influence, even if that influence is mediated by other sociological factors. Regardless of whether the language contact takes place in media contents, in the working experience of people in multilingual media outlets, and/or in the participatory systems enabled by digital social media networks, the importance of media in processes of linguistic contact is clear.

4. References

Alhazmi, Albtool
2015 Linguistic aspects of Arabic-English code-switching on Facebook and radio in Australia. *International Journal of Applied Linguistics and English Literature* 5(3). 184–198.

Amezaga, Josu, Edorta Arana, Patxi Azpillaga & Bea Narbaiza
2012 Minority language television in Europe: Commonalities and differences between regional minority languages and immigrant minority languages. *Zer* 17(32). 89–106.

Androutsopoulos, Jannis
2014a Beyond "media influence." *Journal of Sociolinguistics* 18(2). 242–249.

Androutsopoulos, Jannis (ed.)
2014b *Mediatization and sociolinguistic change*. Berlin: De Gruyter.

Androutsopoulos, Jannis
2014c Mediatization and sociolinguistic change: Key concepts, research traditions, open issues. In Jannis Androutsopoulos (ed.), *Mediatization and sociolinguistic change*, 3–48. Berlin: De Gruyter.

Androutsopoulos, Jannis
2015 Networked multilingualism: Some language practices on Facebook and their implications. *International Journal of Bilingualism* 19(2). 185–205.

Angouri, Jo
2013 The multilingual reality of the multinational workplace: Language policy and language use. *Journal of Multilingual and Multicultural Development* 34(6). 564–581.

Aragrande, Gaia
2016 Investigating multilingual audio-visual and written news: Theoretical and methodological convergence. *Gentes* 3(3). 1–13.
Atkinson, David & Helen Kelly-Holmes
2011 Codeswitching, identity and ownership in Irish radio comedy. *Journal of Pragmatics* 43(1). 251–60.
Bell, Allan
1983 Broadcast news as a language standard. *International Journal of the Sociology of Language* 40. 29–42.
Ben-Rafael, Miriam & Eliezer Ben-Rafael
2012 Plurilingualism in Francophone Comics. In Frank Bramlett (ed.), *Linguistics and the study of comics*, 142–62. London: Palgrave Macmillan.
Bernard-Beziade, Mélanie
2008 L'utilisation des anglicismes dans la langue du sport en Français et leur variation. *Revista de Estudos Linguísticos da Universidade do Porto* 3. 77–94.
Blommaert, Jan
2010 *The sociolinguistics of globalization*. Cambridge: Cambridge University Press.
Cameron, Deborah
2012 *Verbal hygiene*. London & New York: Routledge.
Castells, Manuel
2005 *The network society: A cross-cultural perspective*. Cheltenham: Edward Elgar.
Catenaccio, Paola, Colleen Cotter, Mark De Smedt, Giuliana Garzone, Geert Jacobs, Felicitas Macgilchrist, Lutgard Lams, Daniel Perrin, John E. Richardson, Tom Van Hout & Ellen Van Praet
2011 Towards a linguistics of news production. *Journal of pragmatics* 43(7). 1843–1852.
Coupland, Nikolas
2014 Sociolinguistic change, vernacularization and broadcast British media. In Jannis Androutsopoulos (ed.), *Mediatization and sociolinguistic change*, 67–96. Berlin: De Gruyter.
Cru, Josep
2017 Bilingual rapping in Yucatán, Mexico: Strategic choices for Maya language legitimation and revitalisation. *International Journal of Bilingual Education and Bilingualism* 20(5). 481–496.
Darquennes, Jeroen & Wim Vandenbussche
2015 The standardisation of minority languages: Introductory remarks. *Sociolinguistica* 29(1). 1–16.
de Bres, Julia & Luc Belling
2015 Free your stuff Luxembourg! Language policies, practices and ideologies in a multilingual Facebook group. *Language Policy* 14(4). 357–75.
Díaz Noci, Javier
2002 Lengua estándar, dialectos y medios de comunicación: El Euskera. *Mediatika* 9. 161–83.
Fayzrakhmanova, Yu
2016 Koreanized English words from perspectives of Korean-English language contact. *Asian Englishes* 18(3). 216–231.
García Canclini, Néstor
2001 *Consumers and citizens: Globalization and multicultural conflicts*. Minneapolis: University of Minnesota Press.
Gaudet, Sophie & Richard Clément
2009 Forging an identity as a linguistic minority: Intra- and intergroup aspects of language, communication and identity in western Canada. *International Journal of Intercultural Relations* 33(3). 213–227.

Glynn, Kevin & A. F. Tyson
2007 Indigeneity, media and cultural globalization: The case of Mataku, or the Maori X-Files. *International Journal of Cultural Studies* 10. 205–224. http://dx.doi.org/10.1177/1367877907076788.
Jenkins, Henry, Mizuko Ito & Danah Boyd
2016 *Participatory culture in a networked era: A conversation on youth, learning, commerce and politics.* Cambridge: Polity Press.
Jones, Elin Haf Gruffydd
2013 Permeable and impermeable linguistic boundaries: From mass media to social media in policy and practice in minoritised language contexts. *Zer* 18(35). 29–45.
Kelly-Holmes, Helen, Máiréad Moriarty & Sari Pietikäinen
2009 Convergence and divergence in Basque, Irish and Sámi media language policy. *Language Policy* 8(1). 227–242.
Kelly-Holmes, Helen & Tommaso M. Milani
2011 Thematising multilingualism in the media. *Journal of Language and Politics* 10(4). 467–89.
Lee, Jamie Shinhee
2006 Linguistic constructions of modernity: English mixing in Korean television commercials. *Language in Society* 35(1). 59–91.
Mano, Winston
2006 African Public service radio versus national languages: Mixed responses to Radio Zimbabwe's bilingual service. In Abiodun Salawu (ed.), *Indigenous language media in Africa*, 271–305. Lagos: CBAAC.
McLuhan, Marshall
1994 *Comprender los medios de comunicación: Las extensiones del ser humano*. Barcelona: Paidós.
Meesters, Gert
2012 To and fro Dutch Dutch: Diachronic language variation in Flemish comics. In Frank Bramlett (ed.), *Linguistics and the study of comics*, 163–82. London: Palgrave Macmillan.
Moriarty, Mairead
2009 Normalising language through television: The case of the Irish language television channel, TG4. *Journal of Multicultural Discourses* 4(2). 137–149.
Paffey, Darren
2007 Policing the Spanish language debate: Verbal hygiene and the Spanish language academy (Real Academia Española). *Language Policy* 6(3–4). 313–332.
Patro, Jasabanta, Bidisha Samanta, Saurabh Singh, Prithwish Mukherjee, Monojit Choudhury & Animesh Mukherjee
2017 Is this word borrowed? An automatic approach to quantify the likeliness of borrowing in social media. *arXiv* March. 1–10.
Perrin, Daniel
2017 Medialinguistic approaches: Exploring the case of newswriting. In Colleen Cotter & Daniel Perrin (eds.), *The Routledge handbook of language and media*, 9–26. London: Routledge.
Pietikäinen, Sari & Helen Kelly-Holmes
2011 Gifting, service, and performance: Three eras in minority-language media policy and practice. *International Journal of Applied Linguistics* 21(1). 51–70.
Ramallo, Fernando
2017 Minority languages in media communication. In Kristina Bedijs & Christiane Maas (eds.), *Manual of Romance languages in the media*, 453–470. Berlin: De Gruyter.

Schmidt, Lauren B.
2015 Effects of regional media, travel, and social contacts on the perception of Spanish aspirated-/s/. *Borealis* 4(1). 99–120.
Selnow, Gary W. & Erwin P. Bettinghaus
2009 Television exposure and language development. *Journal of Broadcasting* 26(1). 469–79.
Stuart-Smith, Jane
2007 The influence of the media. In Carmen Llamas, Louise Mullany & Peter Stockwell (eds.), *The Routledge companion to sociolinguistics*, 140–148. London: Routledge.
Stuart-Smith, Jane
2017 Sociolinguistic approaches: Variationist frameworks. In Colleen Cotter & Daniel Perrin (eds.), *The Routledge handbook of language and media*, 27–43. London & New York: Routledge.
Stuart-Smith, Jane, Gwilym Pryce, Claire Timmins & Barrie Gunter
2013 Television can also be a factor in language change: Evidence from an urban dialect. *Language* 89(3). 501–536.
Thomason, Sarah
2013 Contact explanations in linguistics. In Raymond Hickey (ed.), *The handbook of language contact*, 31–47. Malden, MA & Oxford: Wiley-Blackwell.
Tsuda, Sanae & Beverley Elsom Lafaye
2005 Japan's experience of language contact: A case study of RADIO-I, a multilingual radio station in Nagoya. *Language and Intercultural Communication* 5(3–4). 248–263.
Uribe-Jongbloed, Enrique
2016 Issues of identity in minority language media production in Colombia and Wales. *Journal of Multilingual and Multicultural Development* 37(6). 615–627.
Uribe-Jongbloed, Enrique & María Fernanda Peña Sarmiento
2014 Negociaciones de identidad en la radio indígena y étnica Colombiana : Tres casos de estudio. *Anagramas* 13(25). 167–188.
Velázquez, Isabel
2017 Reported literacy, media consumption and social media use as measures of relevance of Spanish as a heritage language. *International Journal of Bilingualism* 21(1). 21–33.
Zhang, Jingning
2009 Mandarin maintenance among immigrant children from the People's Republic of China: An examination of individual networks of linguistic contact. *Language, Culture and Curriculum* 22(3). 195–213.
Zhuravleva, Anastasia, Kees de Bot & Nanna Haug Hilton
2016 Using social media to measure language use. *Journal of Multilingual and Multicultural Development* 37(6). 601–614.
Zipp, Lena
2017 Code-switching in the media: Identity negotiations in a Gujarati diaspora radio program. *International Journal of the Sociology of Language* 247. 33–48.

Enrique Uribe-Jongbloed, Bogotá (Colombia)

36. Language contact in the linguistic landscape

1. Introduction
2. Pioneering studies
3. Early adopters
4. Current research
5. Evaluation and outlook
6. References

1. Introduction

"In linguistic landscapes two or more languages will be said to be in contact if they are used on the same sign". This could be a definition for the field of linguistic landscape studies, in which the study of language(s) on signage in public spaces takes central stage. In actual fact, the statement rephrases the opening sentence of Weinreich's (1953) classic book on *Languages in Contact*: "[i]n the present study, two or more languages will be said to be in contact if they are used alternately by the same persons". In his PhD thesis, on which said book is based, Weinreich ([1951] 2011: 128) demonstrated some awareness of the linguistic landscape, in that he illustrated the importance of language contact by including three pictures of bilingual French-German signage in the city of Biel, Switzerland. Linguistic landscape (LL) studies can thus be seen as a case of contact linguistics.

Although they were not the first to use the concept, Landry and Bourhis (1997) are often quoted when researchers define what LL studies are about. The following two definitions are frequently used (see Gorter [2018] for a broader discussion): (1) "[l]inguistic landscape refers to the visibility and salience of languages on public and commercial signs in a given territory or region" (Landry and Bourhis 1997: 23); and (2) "[t]he language of public road signs, advertising billboards, street names, place names, commercial shop signs, and public signs on government buildings combine to form the linguistic landscape of a given territory, region or urban agglomeration" (Landry and Bourhis 1997: 25). A more up-to-date and concise definition is provided by Shohamy and Ben-Rafael (2015: 1), according to whom the "main goal of LL studies is to describe and identify systematic patterns of the presence and absence of languages in public spaces and to understand the motives, pressures, ideologies, reactions and decision making of people regarding the creation of LL in its varied forms".

The languages as they are displayed in public are studied in various overlapping fields of linguistics, including sociolinguistics, applied linguistics, contact linguistics, language policy, etc. Some of the frameworks (e.g. frame analysis and the advocacy coalition framework as it is applied in language policy research) and concepts (e.g. superdiversity) developed in those fields are applied to the study of LL phenomena. LL research also profits from developments in neighboring disciplines of linguistics, including anthropology, human geography, and urban studies. Obviously, LL studies are not only inspired by other subdisciplines of linguistics; they also (have the potential to) influence research directions in those fields, including in the field of contact linguistics. While Woldemariam and Lanza (2014: 80) state that "the issue of language contact in the linguistic landscape has been rarely addressed", a quick survey of pioneering studies (Section 2) and early adopters (Section 3) shows that quite a number of LL studies do in fact address (albeit societal rather than purely linguistic) issues of language contact.

https://doi.org/10.1515/9783110435351-036

The interest of LL studies in language contact phenomena is even more outspoken in recent work (Section 4), and also promises to color future directions in LL studies (Section 5).

2. Pioneering studies

The field of LL studies has emerged relatively recently, even if the history of LL goes back far in time. In discussing the early functions of writing, Coulmas (2009: 13) claims that "linguistic landscaping is as old as writing", the reason being that writing is also concerned with the public display of inscriptions on monuments, borders, and property markers. Coulmas (2009) discusses the producers, locations, functions, and readers of some examples in antiquity, such as the trilingual *Rosetta Stone*, the bilingual *Behistun Inscription*, and the writing on the wall of the *Menetekel*. He concludes that these ancient inscriptions "embody many of the intricacies of language contact, language choice and linguistic hierarchy that form the substance of linguistic landscape research" (Coulmas 2009: 18). His arguments make clear that language contact is an important element of LL studies. Pointing out that there has always been a strong interplay between processes of urbanization and the development of writing, Coulmas' (2009: 13) work also helps to explain why most LL studies − including the pioneering studies discussed below − tend to concentrate on cities.

One of the earliest pioneering LL studies took place in Tokyo, Japan; this is perhaps surprising in a city and country which is often perceived to be predominately monolingual. Backhaus (2007: 48) summarizes a study of Tokyo's linguistic landscape (*gengo keikan*), based on data collected in 1962 by the geographer Masai, which documents the languages and scripts used on shop signs in the Shinjuku area of the metropolis (see Backhaus [2007: 48−49] for a detailed summary). After describing three additional Japanese LL studies, Backhaus (2007: 53) concludes that "previous research into the linguistic landscape of Tokyo has shown that the streets of the city are a place of written language contact".

Less surprising are the results of some early LL studies in Canada, a country associated with a long history of French-English language contact (and conflict) in the province of Quebec. Early work on the linguistic landscape (*paysage linguistique*) was carried out by Monnier (1989), who surveyed the commercial sector of Montreal. The *Office québécois de la langue française* (Quebec Board of the French Language) later carried out similar inventories of the language(s) used on signs, and continues this work to today (www.oqlf.gouv.qc.ca). Landry and Bourhis (1997) additionally report a study on the linguistic vitality of French in Canada. They reflect on the distinction between informative and symbolic functions of signs, and the differences between private and government signs. Based on a factor analysis of their survey data, they conclude that "the linguistic landscape is a sociolinguistic factor distinct from other types of language contacts in multilingual settings" (Landry and Bourhis 1997: 45).

Language contact also characterizes the city of Jerusalem, where Rosenbaum et al. (1977) carried out a study on the spread of English in contact with Hebrew. One part of their study is an analysis of the use of Hebrew and Roman scripts (the latter almost solely for English) on signs on one busy street of Jerusalem. They show that Hebrew

dominated, yet English had a presence: approximately one third of the signs were balanced bilingual, one third used some English next to Hebrew, and the remainder used only Hebrew. Their outcome reflects the official language policy that in those days supported Hebrew-only signs. At the same time, the study shows that the commercial use of English on signs gave way to a juxtaposition of both languages. The researchers concluded that English was seen relatively more than it was heard (Rosenbaum et al. 1977: 151).

Again in Jerusalem, but some years later, Spolsky and Cooper (1991) included the LL as an important part of their sociolinguistic study of the city. At one point, they show that a detailed analysis of the way in which Hebrew, Arabic, and English are presented on one pair of street signs can give insight into language contact, dominance, and historical development (Spolsky and Cooper 1991: 4–8; see also Spolsky 2009: 26–27). Because the signs on opposite sides of the same street obviously dated from different periods of different rulers of the city, the example demonstrates how the study of the LL can clarify historical changes in contact between languages.

Calvet's (1990) comparative study of the linguistic landscapes in Dakar and Paris shows that the authorities do not take multilingualism in either city into account, because although French is the only official language in both, the LL does take different shapes. In Paris, languages such as French, Arabic, and Chinese seem to be maintained separately, aided by different scripts, whereas in Dakar, one does find different languages and scripts on the same sign, where they coexist and interact (Calvet 1990: 79).

Some pioneering LL studies also took place in Brussels, which was historically a Dutch-speaking city, but later faced massive Dutch-French language shift in the 19th and first half of the 20th centuries. Tulp (1978) investigated language contact between Dutch and French by making an inventory of some 1,200 billboards and large posters placed along the main tramlines of the city. Her findings show that, in the late 1970s, French dominated public space and that the two languages, Dutch and French, were kept separate on the signs. There were no instances of Dutch-French code-switching on the same sign, and thus no 'signs' of real contact. Tulp (1978: 277) describes infrequent use of other languages; she does not explicitly mention the use of English on billboards, which seems to indicate it was not a relevant phenomenon at the time. In a follow-up study that was carried out in 1992, Wenzel (1998) repeated Tulp's study. She found that French still dominated, but also that almost 10 percent of all posters were in English. In a third, more recent study in 2009–2010, Vandenbroucke (2015) found that in some locations, Dutch and English had a similar presence of around 20 percent, even if French remained dominant in Brussels' public space. Other languages, such as Arabic, Chinese, Japanese, Italian, and Spanish can also be observed, although in small numbers. Also relying on the work of many of the early adopters, Vandenbroucke (2015: 175) concludes that "the occurrence of these languages in Brussels' landscapes is a clear indication of transnational mobility and globalisation".

3. Early adopters

As a field proper, LL studies emerged only since 2006. According to Van Mensel, Vandenbroucke, and Blackwood (2016), the first impulse was given by the publication of a

special issue of the *International Journal of Multilingualism*, which was also published as an edited volume (Gorter 2006b). A second influential publication was a monograph about multilingualism in the LL in Tokyo (Backhaus 2007). Laying the groundwork for the field, these volumes were soon followed by two book collections (Jaworski and Thurlow 2010; Shohamy and Gorter 2009), which, along with Scollon and Wong-Scollon (2003), paved the way for the field's rapid expansion. These books, as well as a number of articles (e.g. Coupland and Garrett 2010; Leeman and Modan 2009), make it clear that LL studies vary in their theoretical and methodological approaches. They also demonstrate substantial linguistic and geographic variation. As shown below, divergence remains a characteristic inherent to the analysis of languages in public space.

Parts of the work from the early adopters focus on aspects of language contact. Huebner (2006), for example, explores questions of language contact, mixing, and dominance in different neighborhoods of Bangkok, the capital of Thailand. Although he presents detailed quantitative data on differences between neighborhoods in terms of the distribution and dominance of Thai on signs, his main argument is about language contact, mixing, and change. By analyzing a number of individual examples of signage, he observes a nonreciprocal relationship in which there is a strong influence of the English lexicon and grammar on the development of Thai, including on the Thai script. This leads him to conclude that, on a theoretical level, this calls into question the notion of 'speech community', and "even what constitutes a language itself" (Huebner 2006: 50).

Questions related to the way in which languages coexist, interact, compete, and interfere with each other in certain settings are also addressed in a comparative study of urban multilingualism in Tokyo by Backhaus (2007). The opening sentence of his book states that "[t]he city is a place of language contact" (Backhaus 2007: 1), which shows the importance of this theme in his study. Through his analysis of the LL, he uncovers the multilingual reality of Tokyo, hitherto known as a monolingual city. His book also contains a long and detailed historical overview of the LL field, in which Backhaus demonstrates that earlier studies of LL have tried to offer different perspectives on phenomena of language contact, such as those mentioned in this chapter using data from cities like Montreal, Jerusalem, Bangkok, Brussels, and so on.

In his chapter, Gorter (2006a) sketches further possibilities for LL research. He argues for 'multiple perspectives', including additional research on the link between LL and language policy, but also for incorporating ideas from city planning, from psychological experiments in visual perception, and/or from second language acquisition (cf. Cenoz and Gorter 2008). He also points out how the development of new technology may influence the future direction of the LL field.

Other publications that are part of the early adopters influenced much of the recent work on LL, and implicitly, rather than explicitly, deal with language contact phenomena, including Jaworski and Thurlow (2010: 3), who present a program based on human geography and art history that conceives of landscape "as a way of seeing". Space is not only physical, but also socially constructed. They recommend context-specific analyses of language and encourage a multifaceted, in-depth ethnographic approach. Although their program does not deal with language contact, it can be used as approach for such studies. In a similar vein, Coupland and Garrett (2010: 12) comment that LL research needs to "be more sensitive to historical processes and contexts, as well as to textual nuances, and to move beyond descriptivist and distributional approaches [because] more qualitative approaches, then, can ask different questions". Leeman and Modan (2009) do

apply such a qualitative approach in their analysis of the LL in Chinatown in Washington, DC, which is based on the notions set forth in the framework of geosemiotics and nexus analysis (Scollon and Wong-Scollon 2003).

The studies of the early adopters laid the groundwork for expansion of the field in the past decade. In the next section, we will home in on more recent studies. Given the purpose of this chapter, the overview will concentrate on studies in which issues of language contact play more than just an implicit role. A broader picture can be found in Van Mensel, Vandenbroucke, and Blackwood (2016: 423), who emphasize that LL research has come to be "kaleidoscopic in nature".

4. Current research

Multilingualism has long been a central topic in LL research, and quite a number of recent publications focus on the different ways in which multilingualism manifests itself in written form on signs in − usually public − spaces in a variety of geographical and linguistic settings. As illustrated below, LL research intersects with research on phenomena such as code-mixing, linguistic repertoires, translanguaging, polylanguaging, and metrolingualism.

Kasanga (2010), for example, investigates code-mixing between French (as the local lingua franca) and English (as a more global language) on signs in the Democratic Republic of Congo. Kasanga (2010: 191−197) distinguishes four functions of English in French advertising: brand (i.e. English names of brands to preserve global corporate image), hybrid (i.e. adding French tags to English names), clone (i.e. copying well-known brand names), and imitation (i.e. using English labels in French rather than English brand names in French advertisements). Another example, from a more linguistic point of departure, James (2016) broadens our perspective on language contact by examining how English as the global language mixes with local languages in numerous ways. He presents cases of tourist advertisements to demonstrate that visual English has a graphic impact, for example, in a tourist slogan for the country Slovenia, emphasizing the embedded letters 'LOVE'. He concludes that English in contact with local languages connects and creates local meanings and identities through channel, mode, modality, and medium. The semiotic effect of English, together with local languages, is multimodal and multilingual.

LL research increasingly contributes to discussions that question the existence of languages as bounded entities. Huebner's (2006: 50) analysis of English-Thai mixing led him to challenge the idea of a 'language' at a theoretical level, asking "[w]here […] one language end[s] and the other begin[s]". In their study of billboards and road signs in Delhi, Agnihotri and McCormick (2010: 58) come to the conclusion that the signs they examined "showed varying degrees of language separation and blurring of linguistic boundaries". Woldemariam and Lanza (2014) also investigate signage in Ethiopia. They focus on structural linguistic phenomena of language contact between Amharic, the federal language, and two regional languages, Tigrinya and Oromo. In the text on the signs, the structural linguistic influences of Amharic on both languages are reflected in aspects such as word order, morphological markers, and/or hybrid constructions, which shows how the boundaries between the languages are blurred.

Such ideas are, of course, consistent with the so-called 'multilingual turn' in applied linguistics (May 2014). They also show the interest in contact phenomena in LL studies.

The concept of language as a fixed and bounded system is replaced by ideas about languages (or languaging) as fluid and dynamic resources that are part of the language practices of multilinguals. The notion of 'translanguaging' helps to capture those multilingual practices, and is increasingly applied to the linguistic landscape, as is shown by Gorter and Cenoz (2015). For them, multilingualism (as opposed to separate languages) should be the point of departure for LL studies. The issue is not whether signs are multilingual units, but rather the central question becomes a holistic approach: "[h]ow are multilingual units and multilingual neighborhoods shaped in a social context?" (Gorter and Cenoz 2015: 71). The most appropriate level for assessing and studying multilingualism and translanguaging in linguistic landscapes is the aggregate of signs at the neighborhood level, where "the individual signs combine, alternate and mix to shape linguistic landscapes as whole" (Gorter and Cenoz 2015: 54). The focus then shifts away from contact between separate languages, to language fluidity.

In addition to a focus on more linguistic or 'structural' features of language contact, recent LL studies are also characterized by an interest in the social dimension of multilingualism, as was already present in the work of Landry and Bourhis (1997) and others. When it comes to the social dimension of multilingualism in LL studies, Van Mensel, Marten, and Gorter (2012: 319) argue in a rather programmatic way that an LL approach to the study of minority languages provides a new lens to look "at the 'same old issues' of language contact and language conflict". Not only in the case of minority language settings do such issues include, yet are not limited to: language revitalization (an increase in the public visibility of a minority language contributes to language revitalization; see, for example, the case of Basque in Aiestaran, Cenoz, and Gorter [2010]), language tokenism (the tokenistic use of threatened languages for touristic purposes; see e.g. Hornsby [2008] for Breton, or Moriarty [2012] for Ireland), language conflict (officially and unofficially regulated signs are linked to the language behavior of the population in a given region, which may or may not be conflictual, as shown by Van Mensel and Darquennes [2012] in a study on the German-speaking community in Belgium), intergroup relations and the negotiation of linguistic identities (as shown in Shohamy and Waksman's [2010] study of LL items produced for the centennial of the city of Tel Aviv-Jaffa), and language policy issues in general (and more specifically the tension between 'top-down' policies and LL practice on signs in both the public and private spheres; see e.g. Hassa 2012).

5. Evaluation and outlook

The field of LL studies in general can be conceived of as a subfield of sociolinguistics or applied linguistics. As has become obvious from the foregoing discussion, various research projects have contributed insights to contact linguistics. Mainly urban contexts are investigated as sites of language contact, in which languages are used on the same sign, juxtaposed, mixed, or merged into hybrids.

The direction the field seems to be taking is not directly linked to contact linguistics, and different interpretations exist. For example, Moriarty (2014: 459), Zabrodskaja and Milani (2014: 2), and Lanza and Woldemariam (2015: 177) suggest that LLs studies are moving away from 'documenting signs' or 'counting languages', and in their opinion,

the field has taken a 'qualitative turn'. However, those authors seem to overlook, on the one hand, the qualitative work done by some pioneers and early adopters (e.g. Calvet 1990; Cenoz and Gorter 2008; Reh 2004; Spolsky and Cooper 1991), while also neglecting the qualitative aspects of such earlier work, which usually deals with more than just distributional issues. On the other hand, they also ignore the continuous stream of publications based on quantitative methods (e.g. Coluzzi 2012; Macalister 2012; Neves 2016; Soukup 2016). The direction of LL studies takes further shape through two contrasting proposals for the way to carry out LL work: an Ethnographic Linguistic Landscape Approach (ELLA) and a Variationist Linguistic Landscape Study (VaLLS).

The 'Ethnographic Linguistic Landscape Approach', or ELLA, is elaborated by Blommaert and Maly (2016). The basis for this approach was supplied by Blommaert (2013: 16), who carried out a qualitative analysis of his own neighborhood in Antwerp in which he argues for LL studies as a form of ethnography in order to "bring out its full descriptive and explanatory potential". The ELLA approach is intended to enable researchers to diagnose and provide a detailed image of the demography of a neighborhood, not only in terms of distributions, but one which also consists of several layers. Distinctions can be made between older and newer groups, and how practices and relationships between groups are organized. One of the ideas is that the ambitions and identity aspirations of different groups can be shown through elements of signage, which are all part of historical processes of transformation. The same themes are also central in a case study by Blommaert and Maly (2016), who studied public signage in a highly diverse neighborhood in the Belgium city of Ghent. In their analysis of the signage, they emphasized the historically layered view of multilingualism and multiculturalism. Although language contact is not an explicit focus of this approach, it has relevance for the study of contact between languages and cultures, and for related issues about social change in complex, superdiverse neighborhoods.

A contrasting approach is called VaLLS, an acronym for 'Variationist Linguistic Landscape Study'. Soukup (2016) describes how the principles and standards of VaLLS came from variationist sociolinguistics (which is typically quantitative in nature), and she applies the approach to linguistic variation on signs in her study of Vienna. English plays a role in language contact situations around the world, and in her proposed study, Soukup plans to analyze correlations between social factors and the choice of English on signs.

The two proposals are too recent to know if either approach − ELLA or VaLLS − will be adopted by others and become an important influence on LL studies. It seems, however, more likely that the field will continue to include a rich diversity of qualitative, quantitative, and mixed method approaches, in which some studies will emphasize language contact phenomena more than others.

Just by looking at the number of publications, one can observe that the LL field has developed exponentially. Previous research published through 2006 was summarized by Backhaus (2007), and at that time, he could list only 30 publications, of which 10 were from prior to 1998. Some years later, Troyer (2014) presented his online bibliography of LL publications. His bibliography at that moment contained 287 publications, of which no fewer than 235 publications had appeared between 2007 and 2014. As of December 2018, the bibliography has more than doubled and now contains over 700 entries.

It is obvious that language contact will continue to play an inevitable role in LL research. After all, while LL studies may be relevant in monolingual contexts, investigations of multilingual linguistic landscapes are far more exciting when they deal with variation in, conflict among, and contact of multiple languages.

Acknowledgements

This work was supported by the MINECO/FEDER [Grant number EDU2015–63967-R] and the Basque Government [Grant number DREAM IT-714–13; UFI 11/54].

6. References

Agnihotri, Rama Kant & Kay McCormick
2010 Language in the material world: Multilinguality in signage. *International Multilingual Research Journal* 4(1). 55–81.

Aiestaran, Jokin, Jasone Cenoz & Durk Gorter
2010 Multilingual cityscapes: Perceptions and preferences of the inhabitants of Donostia-San Sebastián. In Elana Shohamy, Eliezer Ben-Rafael & Monica Barni (eds.) *Linguistic landscape in the city*, 219–234. Bristol: Multilingual Matters.

Backhaus, Peter
2007 *Linguistic landscapes: A comparative study of urban multilingualism in Tokyo*. Clevedon: Multilingual Matters.

Blommaert, Jan
2013 *Ethnography, superdiversity, and linguistic landscapes: Chronicles of complexity*. Bristol: Multilingual Matters.

Blommaert, Jan & Ico Maly
2016 Ethnographic linguistic landscape analysis and social change: A case study. In Jan Blommaert, Ben Rampton, Karel Arnaut and Massimiliano Spotti (eds.), *Language and superdiversity*, 191–211. London: Routledge.

Calvet, Louis Jean
1990 Des mots sur les murs: Une comparaison entre Paris et Dakar. In Robert Chaudenson (ed.), *Des langues et des villes: Actes du colloque international à Dakar, du 15 au 17 décembre 1990*, 73–83. Paris: Didier Érudition.

Cenoz, Jasone & Durk Gorter
2008 Linguistic landscape as an additional source of input in second language acquisition. *IRAL: International Review of Applied Linguistics in Language Teaching* 46(3). 267–287.

Coluzzi, Paolo
2012 Multilingual societies vs. monolingual states: The linguistic landscape of Italy and Brunei Darussalam. In Durk Gorter, Luk Van Mensel & Heiko F. Marten (eds.), *Minority languages in the linguistic landscape*, 225–242. Basingstoke: Palgrave Macmillan.

Coulmas, Florian
2009 Linguistic landscaping and the seed of the public sphere. In Elana Shohamy & Durk Gorter (eds.), *Linguistic landscape: Expanding the scenery*, 13–24. London: Routledge.

Coupland, Nik & Peter Garrett
2010 Linguistic landscapes, discursive frames and metacultural performance: The case of Welsh Patagonia. *International Journal of the Sociology of Language* 205. 7–36.

Gorter, Durk
2006a Further possibilities for linguistic landscape research. In Durk Gorter (ed.), *Linguistic landscape: A new approach to multilingualism*, 81–89. Clevedon: Multilingual Matters.
Gorter, Durk (ed.)
2006b *Linguistic landscape: A new approach to multilingualism*. Clevedon: Multilingual Matters.
Gorter, Durk
2019 Methods and techniques for linguistic landscape research: About definitions, core issues and technological innovations. In Martin Pütz & Neele Mundt (eds.), *Expanding the linguistic landscape: Multilingualism, language policy and the use of space as a semiotic resource*, 38–57. Bristol: Multilingual Matters.
Gorter, Durk & Jasone Cenoz
2015 Translanguaging and linguistic landscapes. *Linguistic Landscape* 1(1). 54–74.
Hassa, Samira
2012 Regulating and negotiating linguistic diversity: Top-down and bottom-up language planning in the Moroccan city. *Current Issues in Language Planning* 13(3). 207–223.
Hornsby, Michael
2008 The incongruence of the Breton linguistic landscape for young speakers of Breton. *Journal of Multilingual and Multicultural Development* 29(2). 127–138.
Huebner, Thom
2006 Bangkok's linguistic landscapes: Environmental print, codemixing, and language change. In Durk Gorter (ed.), *Linguistic landscape: A new approach to multilingualism*, 31–51. Clevedon: Multilingual Matters.
Huebner, Thom
2016 Linguistic landscape: History, trajectory and pedagogy. *Manusya* 22. 1–11.
James, Allen
2016 From code-mixing to mode-mixing in the European context. *World Englishes* 35(2). 259–275.
Jaworski, Adam & Crispin Thurlow (eds.)
2010 *Semiotic landscapes: Language, image, space*. London: Continuum.
Kasanga, Luanga A.
2010 Streetwise English and French advertising in multilingual D. R. Congo: Symbolism, modernity, and cosmopolitanism. *International Journal of the Sociology of Language* 206. 181–205.
Landry, Roger & Richard Y. Bourhis
1997 Linguistic landscape and ethnolinguistic vitality: An empirical study. *Journal of Language and Social Psychology* 16(1). 23–49.
Lanza, Elizabeth & Hirut Woldemariam
2015 Imagined community: The linguistic landscape in a diaspora. *Linguistic Landscape: An International Journal* 1(1/2). 172–190.
Leeman, Jennifer & Gabrielle Modan
2009 Commodified language in Chinatown: A contextualized approach to linguistic landscape. *Journal of Sociolinguistics* 13(3). 332–362.
Macalister, John
2012 Language policies, language planning and linguistic landscapes in Timor-Leste. *Language Problems & Language Planning* 36(1). 25–45.
May, Stephen (ed.)
2014 *The multilingual turn: Implications for SLA, TESOL and bilingual education*. New York: Routledge.
Monnier, Daniel
1989 *Langue d'accueil et langue de service dans les commerces à Montréal*. Québec, Canada: Conseil de la langue française.

Moriarty, Máiréad
2012 Language Ideological Debates in the Linguistic Landscape of an Irish Tourist Town. In Durk Gorter, Heiko F. Marten, & Luk Van Mensel (eds.), *Minority languages in the linguistic landscape*, 74–88. Basingstoke: Palgrave-Macmillan.
Moriarty, Máiréad
2014 Contesting language ideologies in the linguistic landscape of an Irish tourist town. *International Journal of Bilingualism* 18(5). 456–463.
Neves, Ana. C.
2016 Linguistic landscape of Macau: A quantitative analysis. In Lydia Sciriha (ed.), *International perspectives on bilingualism*, 43–62. Newcastle upon Tyne, UK: Cambridge Scholars Publishing.
Reh, Mechtild
2004 Multilingual writing: A reader-oriented typology – with examples from Lira Municipality (Uganda). *International Journal Sociology of Language* 170. 1–41.
Rosenbaum, Yehudit, Elizabeth Nadel, Robert L. Cooper & Joshua A. Fishman
1977 English on Keren Kayemet Street. In Joshua A. Fishman, Robert L. Cooper & Andrew W. Conrad (eds.), *The spread of English: The sociology of English as an additional language*, 179–196. Rowley, MA: Newbury House.
Scollon, Ron & Suzie Wong-Scollon
2003 *Discourses in place*. London: Routledge.
Shohamy, Elana & Durk Gorter (eds.)
2009 *Linguistic landscape: Expanding the scenery.* London: Routledge.
Shohamy, Elana & Eliezer Ben-Rafael
2015 Introduction: Linguistic Landscape, a new journal. *Linguistic Landscape* 1(1/2). 1–5.
Shohamy, Elana & Soshi Waksman
2010 Decorating the city of Tel Aviv-Jaffa for its centennial complementary narratives via linguistic landscape. In Elana Shohamy, Eliezer Ben-Rafael & Monica Barni (eds.), *Linguistic landscape in the city*, 57–73. Bristol: Multilingual Matters.
Soukup, Barbara
2016 English in the linguistic landscape of Vienna, Austria (ELLViA): Outline, rationale, and methodology of a large-scale empirical project on language choice on public signs from the perspective of sign-readers. *Vienna English Working Papers* 25 (June). 1–25.
Spolsky, Bernard
2009 Prolegomena to a sociolinguistic theory of public signage. In Elana Shohamy & Durk Gorter (eds.), *Linguistic landscape: Expanding the scenery*, 13–24. London: Routledge.
Spolsky, Bernhard & Robert L. Cooper
1991 *The languages of Jerusalem*. Oxford: Clarendon.
Troyer, Robert
2014 Historical and emerging trends in linguistic landscape studies. Poster presented at annual conference American Association of Applied Linguistics (AAAL), Portland, 22–25 March.
Tulp, Stella M.
1978 Reklame en tweetaligheid: Een onderzoek naar de geographische verspreiding van franstalige en nederlandstalige affiches in Brussel. *Taal en Sociale Integratie* 1. 261–288.
Vandenbroucke, Mieke
2015 Language visibility, functionality and meaning across various TimeSpace scales in Brussels' multilingual landscapes. *Journal of Multilingual and Multicultural Development* 36(2). 163–181.
Van Mensel, Luk & Jeroen Darquennes
2012 All is quiet on the Eastern Front? Language contact along the French- German language border in Belgium. In Durk Gorter, Heiko F. Marten, & Luk Van Mensel (eds.), *Minority languages in the linguistic landscape*, 164–180. Basingstroke: Palgrave-Macmillan.

Van Mensel, Luk, Heiko F. Marten & Durk Gorter
2012 Minority languages through the lens of the linguistic landscape. In Durk Gorter, Heiko F. Marten & Luk Van Mensel (eds.), *Minority languages in the linguistic landscape*, 319–323. Basingstoke: Palgrave Macmillan.
Van Mensel, Luk, Mieke Vandenbroucke & Robert Blackwood
2016 Linguistic landscapes. In Ofelia García, Nelson Flores & Massimiliano Spotti (eds.), *Oxford handbook of language and society*, 423–449. Oxford: Oxford University Press.
Weinreich, Uriel
1953 *Languages in contact: Findings and problems*. The Hague: Mouton Publishers.
Weinreich, Uriel
2011 [1951] *Languages in contact: French, German and Romansh in twentieth-century Switzerland*, reprint of dissertation. Amsterdam: John Benjamins Publishing.
Wenzel, Veronika
1998 Reclame en tweetaligheid in Brussel: Een empirisch onderzoek naar de spreiding van Nederlandstalige en en Franstalige affiches. In Piet Van De Craen (ed.), *Mondig Brussel*, 45–73. Brussel: Brio.
Woldemariam, Hirut & Elizabeth Lanza
2014 Language contact, agency and power in the linguistic landscape of two regional capitals of Ethiopia. *International Journal of the Sociology of Language* 228. 79–104.
Zabrodskaja, Anastassia & Tomasso M. Milani
2014 Signs in context: Multilingual and multimodal texts in semiotic space. *International Journal of the Sociology of Language* 228. 1–6.

Durk Gorter, Donostia/San Sebastián (Spain)

37. Language shift

1. Introduction
2. Language shift: definition and causes
3. Macro-level investigations
4. Micro-level investigations
5. Linguistic outcomes of language shift
6. Conclusion
7. References

1. Introduction

This chapter provides an overview of the study of language shift. After an introduction to some definitions and main causes of language shift, it broadly divides research on language shift into two major areas: one that takes a macrosociological approach to the investigation of the causes of shift (Fishman 1966; Haugen 1953; Kloss 1966), and another that focuses on the social meaning of language varieties in multilingual speech communities and the symbolic messages associated with language choice (Gal 1979; Gumperz 1982; Le Page and Tabouret-Keller 1985). The first approach includes more emphasis on institutional and demographic factors, and informs measurements of language vitality and language policy efforts to protect linguistic diversity. While the second

https://doi.org/10.1515/9783110435351-037

approach also considers those factors, its focus of investigation is the social meaning of language choice, particularly in the context of social change. The structural effects of shift on the languages themselves will also be discussed, particularly with respect to the question of whether there are linguistic outcomes that can be specifically attributed to language shift, and not to language contact more generally.

2. Language shift: definition and causes

Language shift is, broadly speaking, the process by which a language user, or (parts of) a speech community, shift to a new primary language of communication in a situation of language contact. Although any type of language contact scenario may in theory result in language shift, Lüpke (2015, citing Vigouroux and Mufwene [2008: 14]) has argued that in African migratory language ecologies with longer traditions of egalitarian societal multilingualism, speakers experience less pressure to give up heritage languages when acquiring languages of wider communication and thus language shift is less common. Conversely, two major social processes that result in multilingual communities with largely unequal power distributions that regularly contribute to language shift are immigration, and conquest.

Immigration results in a minority language group forming part of a larger host community language group; conquest results in a conquering minority imposing their language on the majority host population, sometimes by force, but often through practices such as universal education (Sankoff 2001; Winford 2003: 237). In no case is shift to exclusive use of a single language necessary in any functional sense: there is no clear benefit of monolingualism to the individual; there are clear benefits of multilingualism (Bialystock 2011); and there is no reason to believe that human societies cannot maintain multilingualism (McLaughlin 2008; Sorensen 1967). Despite this, language shift often occurs in immigration scenarios within two generations, and even in conquest scenarios, where a period of stable multilingualism may last longer, shift to the socially dominant language at the expense of others seems to be the norm. The explanations for language shift, then, are not functionally motivated, but socially motivated.

Due to the nature of the social processes that bring about language shift scenarios, many modern cases involve shift toward global languages of wider communication (LWC; Fishman, Cooper, and Conrad 1977), which are not coincidentally those of former colonial powers, such as English, French, and Spanish. More local LWCs, such as Hindi in India, Hausa in Nigeria, Indonesian in Indonesia (as well as English in the United Kingdom, French in France, Spanish in Spain, etc.) are also implicated in many of the language shift scenarios globally.

Weinreich (1953: 68) provides an early working definition of language shift as a "change from the habitual use of one language to that of another". Fishman (1964: 35) defines it further as "the relationship between change […] in habitual language use […] and ongoing psychological, social, or cultural processes". Clyne (2003: 20) refines these definitions, noting that language shift can be any combination of: i) the language behavior of a whole community, subgroup, or individual; ii) the process (a gradual shifting) or the outcome (language abandonment); and/or iii) a change in the main language, the dominant language, the language of one or more domains, or the exclusive language for

1–3 language skills (listening, speaking, reading, and/or writing). This expanded definition is also a reflection of the different approaches to study of language shift, which are broadly categorized here into macro-level and micro-level investigations.

3. Macro-level investigations

Based largely on the size of the speaker population (under 100,000) as a predictor, Krauss (1992) projected that as many as 90 % of the world's languages were likely to be moribund by the end of the 20th century. However, while a small population size may be a symptom of language shift, it is not clearly a predictor of shift. As early as 1966, Kloss considered size of the speaker population to be ambivalent in terms of its relative importance (Kloss 1966). Recent studies have further demonstrated that size alone does not predict language shift (Abtahian, Cohn, and Pepinsky 2016; Anderbeck 2015; Barreña et al. 2007), as languages with large speaker populations are also affected. Moreover, there is a methodological benefit to focusing on large language communities undergoing shift. With the increasing availability of big data sources, studying large language communities with demographically diverse populations allows researchers to better investigate the interplay of many potentially relevant social factors. Although questionnaires and surveys are also used, census data is often used for this type of analysis.

3.1. Census data

The use of census data to examine language shift is longstanding (Mackey and Cartwright 1979), though not without drawbacks. Problems with the use of census data largely stem from two sources: (1) problems with question text, and (2) inconsistencies in how people respond to the questions. With respect to (1), questions may be formulated in such a way that multilingualism is not captured (i.e. by only leaving a space for one answer, or by asking which language the respondent 'regularly' or 'mainly' uses), or may only allow for restricted answers (which may not include the name(s) a speaker uses for their variety). Moreover, relevant census questions may change from one census to the next, making comparisons unreliable (Clyne 2003: 21–23). With respect to (2), even when the questions are consistent and appropriate, there are a number of reasons why responses may not be. First, as individuals age, their language use patterns may change. Thus, age-related differences that are apparent in a single survey may reflect differences between age groups, but not community change over time (Lieberson 1972, 1980). Moreover, speakers may 'change' their answers as a reflection of the changing sociolinguistic context of the census. Fishman, Gertner, and Lowy (1985) demonstrate this in the US, where the general population rose by 13 percent between the 1960 and 1970 censuses, but the percent of speakers claiming a non-English mother tongue rose by 70 percent. They attribute this change not to linguistic change over time, but rather to the formation of stronger ethnic identities. Conversely, in some social and political climates, speakers may be ashamed, or in some cases afraid, to report their primary language. Darquennes (2007: 350) provides the example of the 1930 and 1947 Belgian censuses, which show a sharp decrease in the number of respondents claiming knowl-

edge of German in German-speaking areas of Belgium, as Belgians no longer wanted to be overtly associated with 'Germanness' in the aftermath of World War II (Darquennes 2007: 351).

A broader problem with census data may be that censuses are neither objective nor apolitical (Arel 2001). There are implicit ideologies reflected in the way questions are asked, and respondents are expected to answer. A census that asks for one's 'mother tongue', for example, presumes that all language varieties are named and considered 'languages' by speakers, and that speakers can name a single variety in which they are most proficient, none of which may reflect actual linguistic practices or beliefs in a community.

One advantage of census data is that it provides us with a rich source of sociodemographic data which can be analyzed alongside linguistic data to provide better representations of the complex nature of the correlations between social factors and language shift, and a longitudinal picture of macro-trends in language use. This in turn can provide necessary information for more accurate measures of language vitality, as well as better data to potentially model language shift, as shown in the next section.

3.2. Social factors, measures of language vitality, and modeling language shift

The study of language shift has from the beginning included an emphasis on identifying causal factors in order to create a typology of language shift. Fishman (1965) focused on the domains of language use as a metric of language shift, and Heinz Kloss' (1966) paper on German-American language maintenance efforts is often cited as the first analysis of the macrosociological factors that correlate with language maintenance. Although some of Kloss' factors are specific to the North American context and the time period, others are more broadly generalizable. Haugen's (1953) study of Norwegian-Americans and his model of language ecology (Haugen 1972) are also early models of this type of approach. Conklin and Laurie (1983) and Edwards (1992) present detailed taxonomies of political, social, demographic, cultural, and linguistic factors that are comprehensive and generalizable, and reflect the complexity of language shift scenarios (see also Himmelman 2010). Out of descriptive necessity, all of these models also include factors that are more difficult to quantify, like 'cohesiveness of the language community', an issue that Bourhis, Giles, and Rosenthal (1981) explicitly try to address with a 'subjective vitality questionnaire'. More recently, the focus in this vein has turned to metrics of language vitality and language policy, including: Giles, Bourhis, and Taylor's (1977) three pillars of ethnolinguistic vitality, namely status, demography, and institutional support; Fishman's (1991) Graded Intergenerational Disruption Scale (GIDS); Lewis and Simons' (2010) Expanded Graded Intergenerational Disruption Scale (EGIDS); the UNESCO scale of endangerment (Brenzinger et al. 2003); and work on the status of minority languages in the European Union (Dell'Aquila and Iannàccaro 2006; Grin et al. 2002; Nelde, Strubell, and Williams 1996; Ó Giollagáin et al. 2007). More recent attempts to model language shift scenarios (Abtahian, Cohn, and Pepinsky 2016; Kandler, Unger, and Steele 2010) also depend largely on macrosocial factors reported in census data and/or large-scale sociolinguistic surveys, and take advantage of statistical models and multivariate analyses.

However, as Gal (1996: 588) points out, generalizations can be made about macrosociological factors, yet none of these factors will apply to every case of language shift; she argues that what is needed is "not a more complex combination of factors, but rather attention to the way that large social changes affect speakers' evaluation of languages and the social, cultural meanings associated with them". In line with this thinking, and underscored by the emphasis placed by Krauss (1992) and others on the need to document languages with small speaker populations, much of the research focus on language shift has correspondingly shifted to the use of ethnographic methods in small speech communities, including at the discourse level of multilingual interactions (Gumperz 1982; Le Page and Tabouret-Keller 1985). In the following section, I present this approach as crucial for understanding certain aspects of language shift, while continuing to recognize the importance of developing comprehensive treatments of the relative prevalence of language shift across communities, for which macrosociological investigations may prove more useful.

4. Micro-level investigations

The 19th century ideology of the nation-state has been accompanied by linguistic ideologies of monolingualism (Blommaert 2006; Clyne 2005; Heller 2007) and standardization (Milroy and Milroy [1985] 2012) that promote the association of one political unit with one linguistic and cultural unit ('imagined communities'; Anderson [1983] 2006). One consequence of this may be the explicit and implicit marginalization of minority languages (Farr and Song 2011; Tollefson 1991). These ideologies are reflected in the language behavior of even individual members of a speech community. Studies of language shift that emphasize the central importance of language ideologies and attitudes generally focus on close examinations of the social meaning of different language varieties in multilingual speakers' repertoires. Dorian's (1981) longitudinal study of a community of East Sutherland Gaelic speakers in Scotland, Gal's (1979) study of a German-Hungarian speaking community in eastern Austria, Kulick's (1992) study of shift toward Tok Pisin in a Papua New Guinea village, and Hill's (1983) work on the shifting role of Nahuatl *vis-à-vis* Spanish are foundational studies in this area. Moreover, while the focus of these studies is typically not on creating models of shift or making generalizations across communities, some generalizations may still be drawn. For example, Gal (1996: 589) points out that a common feature of language shift worldwide seems to be the shifting social symbolism of the use of different codes as a type of response to social change, and the "ultimate association of one code with solidarity and later with a stigmatized local identity".

4.1. Social networks

A productive area of research is the role of social networks (Milroy and Milroy 1978) in determining language choice. In some cases, what at first appears to be a result of social categories like gender (Gal 1979; Zentella 1997) or age/generation (Li Wei 1994) may actually be attributed to social network. Zentella's (1997) study of Spanish-English

use in Harlem, New York, for instance, showed that boys or girls whose language proficiency was dissimilar to their 'gender-mates' were those whose networks were also different from the norm, e.g. the teen male who shared more interests with the girls and spent more time at home was more proficient in Spanish than his "male-bonded, baseball- and football-playing brother" (Zentella 1997: 52). Gal (1979) similarly found that the composition of an individual's network (made up of more 'peasants' or 'workers') correlated with their language choices, and was the strongest predictor of whether they continued to use Hungarian or had shifted to German. In Li Wei's (1994) study of code-switching in a Chinese immigrant community in Tyneside, England, he found that membership in a particular church community − a dense network of speakers of multiple generations − was linked to more maintenance of Chinese for third generation speakers than for non-church-network third generation speakers, who primarily used monolingual English discourse.

4.2. Language ideologies and language attitudes

Equally as important as social network is the social meaning associated with use of a particular language, which implies a close relationship between language attitudes and language shift (Bonner 2001; Garrett, Coupland, and Williams 2003; Gibbons and Ramirez 2004). As Gómez de Garcia, Axelrod, and Lachler (2009: 99) write: "[a]ll bilinguals have attitudes toward the two languages they speak". While a small speaker population or language use in fewer domains may be symptomatic of language shift, attitudes are more causal than symptomatic, and are generally reflective of broader language ideologies. They "interact with governmental policy and societal pressures to result in increased or decreased language use in different domains" (Brenzinger et al. 2003: 15), and as Grenoble and Whaley (1998: 24) argue, "are paramount for predicting shift".

Methods for investigating attitudes are problematic, however, in that they require at least a two-step analysis which first considers whether the data collected is actually reflective of speaker attitudes. The two primary approaches are to explicitly ask speakers about their language beliefs through the use of questionnaires or surveys (e.g. Coronel-Molina 2009; Hoffman and Walker 2010; Vaish 2007), or to try to determine speaker attitudes surreptitiously, as with variations on the matched-guise experiment (Lambert et al. 1960), wherein listeners hear what they are told are different speakers, but is actually one speaker using different language varieties. In the analysis, listeners' judgments are presumed to be representative of their attitudes toward the linguistic variety and/or speakers of it. In many studies, these are triangulated with observations about language use and/or ethnographies of the community with respect to language ideologies.

Social categories, like gender, religion, and education level, that correlate with language shift often do so as a reflection of more complex social meanings associated with language varieties. For example, work on language shift in Indonesia has shown that religion is an important local social factor in language choice (Musgrave and Ewing 2006), not because of religion *per se*, but because, for example, Muslims in Central Java place greater importance on marrying another Muslim than on marrying another Javanese, so speaking Indonesian consequently offers greater social mobility for young educated Muslim-Javanese women (Smith-Hefner 2009). Similarly, Cavanaugh (2006) demonstrates that the role that gender plays in language shift in a northern Italian town is

linked to local language ideologies. In Belize, Abtahian (2017) provides an example of a change in the social meaning of Belizean Creole and Garifuna with respect to changes in the national discourse about the importance of language maintenance and of language as a symbol of culture. Kroskrity and Field (2009) present a number of case studies that similarly link language ideologies and language practices. A growing literature (Childs, Good, and Mitchell 2014; Hill 2002; McCarty, Romero-Little, and Zepeda 2006; Odango 2015) also addresses discourses of language endangerment, and the relationship between linguists' and speakers' view of language shift.

5. Linguistic outcomes of language shift

In most cases, the structural effects of language shift are not due to language shift *per se*, but to the period of language contact that precedes it; most of the linguistic outcomes of shift are covered in separate chapters in this volume (see Part I on linguistic aspects of language contact). However, language shift is especially relevant to certain types of language contact effects, including possible substratum effects of the erstwhile L1 on the newly dominant language, and the possible effects of obsolescence on the erstwhile L1. These two categories are discussed below in the context of the two major social processes introduced earlier: immigration and conquest.

5.1. Linguistic effects of shift on the majority or dominant language

In the context of immigration, a few studies in the United States have reported conservative tendencies in the possible substrate influence on the phonology of some minority ethnic groups (Sankoff 2001: 7), although generally this does not persist for more than a generation. Others have described substratum influence that has persisted longer, for example, in the English of descendants of Yiddish speakers (Prince 1988).

In the context of colonialism, where a longer period of contact may persist before shift occurs to the dominant language, and where the initial majority of speakers are speakers of the non-dominant language, it seems logical that there may be more contact effects on the (minority) dominant language, resulting in what are sometimes termed *indigenous varieties*, such as, for example, World Englishes (Kachru 1985) like Nigerian English and Indian English; Irish English is also given as an example of substantial substratum influence on an incoming, socially-dominant, minority language (Hickey 2010; Winford 2003: 237–242).

In all of these cases, as in others where change is attributed to contact, we should be careful to stress that differences between e.g. Indian English and British English may not be due to contact effects with other Indian languages *per se*, but rather due to ongoing internal change in both varieties. This is also true for changes attributed to the process of obsolescence.

5.2. Linguistic effects of obsolescence

The most common explanation for language change in the context of shift is that the existence or absence of a contrast in the dominant language may promote a change in

the obsolescing language (that may or may not have been already ongoing). Palosaari and Campbell (2011) describe a merger of velar and uvular consonants in the Mam language of Tuxtla Chico in contact with Spanish, which lacks that distinction; Silva-Corvalán (1994) describes syntactic changes in Los Angeles Spanish by analogy with English; Clyne (2003) uses a variety of examples from German, Dutch, and Hungarian immigrants to Australia; Aikhenvald (2002) provides examples from Tariana in contact with newly dominant Tucano in Amazonia.

An additional question exists as to whether there are structural effects of the process of shift itself, i.e. whether, as the language is used in fewer domains or by fewer fluent speakers, there are structural effects on the language itself (Sasse 1992). One possibility is that, as the language is used by a greater number of *semi-speakers*, they change the language (Dorian 1977; Tsitsipis 1989). In this case, structures may be simplified, not necessarily by analogy with simpler structures in the dominant language, but due to the overgeneralization of unmarked features. Cook (1989) describes this in Tuxtla Chico Mam, where marked uvular stops are replaced by unmarked velar stops, and in Chipewyan, where glottalized consonants are replaced by plain consonants. Another possibility is that speakers may overgeneralize marked features that seem 'exotic' (Palosaari and Campbell 2011), a similar process with a nonetheless different linguistic outcome. A related possibility is that, as the language is used in fewer domains or with fewer interlocutors, the social meaning attached to variation in the language diminishes or disappears entirely. This disconnect between sociolinguistic meaning and structural variation may then result in either increased variation in the language (Dressler 1988), or decreased stylistic variation (Dorian 2010; Hill 1978). The nature of the relationship between small language communities, contracting languages, and language variation remains an open area of inquiry.

6. Conclusion

In sum, language shift is both a societal process and a linguistic one, and we gain insight into the process of language shift from different types of explorations. The exploration of macro-factors that correlate with shift, using sources such as censuses and surveys, provides insight into the institutional and demographic factors that correlate with shift, which may help us to create better predictive models of shift and better measures of linguistic vitality and thus to provide support for language planning and policies intended to promote maintenance and revitalization. At a closer level, we can explore individual social networks, multilingual language use in discourse, and the social meaning of different language varieties using ethnographic methods as well as surveys and questionnaires. By situating these in their sociohistorical context, we gain insight into the relationship between large social changes and the symbolic meaning of language choices for individuals. Finally, with careful documentation of both the social context of language shift and the attendant linguistic outcomes, we may gain insight into the structural outcomes of shift and obsolescence.

7. References

Abtahian, Maya R.
2017 Language shift, endangerment and prestige: Kriol and Garifuna in Hopkins, Belize. *Journal of Pidgin and Creole Languages* 32(2). 339–364.

Abtahian, Maya R., Abigail C. Cohn & Thomas Pepinsky
2016 Modeling social factors in language shift. *International Journal of the Sociology of Language* 242. 139–170.

Aikhenvald, Alexandra
2002 *Language contact in Amazonia.* Oxford: Oxford University Press.

Anderbeck, Karl
2015 Portraits of language vitality in the languages of Indonesia. In I Wayan Arka, Ni LuhNyoman Seri Malini & Ida Ayu Made Puspani (eds.), *Language documentation and cultural practices in the Austronesian world*, Papers from ICAL 12 (A-PL 019) 4.19–47. Canberra: Pacific Linguistics.

Anderson, Benedict
2006 [1983] *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.

Arel, Dominique
2001 Language categories in censuses: Backward- or forward-looking? In David I. Kertzer & Dominique Arel (eds.), *Census and identity: The politics of race, ethnicity, and language in national censuses*, 92–120. Cambridge: Cambridge University Press.

Barreña, Andoni, Esti Amorrortu, Ane Ortega, Belen Uranga, Esti Izagirre & Itziar Idiazabal
2007 Does the number of speakers of a language determine its fate? *International Journal of the Sociology of Language* 186. 125–139.

Bialystock, Ellen
2011 Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology* 65(4). 229–235.

Blommaert, Jan
2006 Language policy and national identity. In Thomas Ricento (ed.), *An introduction to language policy, theory and method*, 238–54. Malden, MA: Blackwell.

Bonner, Donna
2001 Garifuna children's language shame: Ethnic stereotypes, national affiliation and transnational immigration as factors in language choice in southern Belize. *Language in Society* 30(1). 81–96.

Bourhis, Richard Y., Howard Giles & Doreen Rosenthal
1981 Notes on the construction of a "subjective vitality questionnaire" for ethnolinguistic groups. *Journal of Multilingual and Multicultural Development* 2(2). 145–155.

Brenzinger, Matthias, Arienne M. Dwyer, Tjeerd de Graaf, Colette Grinevald, Michael Krauss, Osahito Miyaoka, Nicholas Ostler, Osamu Sakiyama, María E. Villalón, Akira Y. Yamamoto & Ofelia Zepeda
2003 *Language vitality and endangerment*. Paris: UNESCO.

Cavanaugh, Jillian R.
2006 Little women and vital champions: Gendered language shift in a northern Italian town. *Journal of Linguistic Anthropology* 16(2). 194–210.

Childs, Tucker, Jeff Good & Alice Mitchell
2014 Beyond the ancestral code: Towards a model for sociolinguistic language documentation. *Language Documentation and Conservation* 8. 168–191.

Clyne, Michael
2003 *Dynamics of language contact: English and immigrant languages*. Cambridge: Cambridge University Press.

Clyne, Michael
2005 *Australia's language potential*. Sydney: USNW Press.
Conklin, Nancy R. & Margaret A. Lourie
1983 *A host of tongues: Language communities in the United States.* New York: Free Press.
Cook, Eung Do
1989 Is phonology going haywire in dying languages? Phonological variations in Chipeewyan and Sarcee. *Language in Society* 18. 235–255.
Coronel-Molina, Serafín M.
2009 Definitions and critical literature review of language attitude, language choice and language shift: Samples of language attitude surveys. https://scholarworks.iu.edu/dspace/handle/2022/3785 (accessed 10 August 2018).
Darquennes, Jeroen
2007 Flirting at the fringe: The status of the German varieties as perceived by language activists in Belgium's Areler Land. In Stephan Elspass, Nils Langer, Joachim Scharloth & Wim Vandenbussche (eds.), *Germanic language histories 'from below' (1700–2000)*, 343–361. Berlin: de Gruyter.
Dell'Aquila, Vittorio & Gabriele Iannàccaro
2006 *Survey Ladins: Usi linguistici nelle Valli Ladine*. Trento: Regione Autonoma Trentino-Alto Adige.
Dorian, Nancy C.
1977 The problem of the semi-speaker in language death. *International Journal of the Sociology of Language* 12. 23–32.
Dorian, Nancy C.
1981 *The life cycle of a Scottish Gaelic dialect.* Philadelphia: University of Pennsylvania Press.
Dorian, Nancy C.
2010 *Investigating variation: The effects of social organization and social setting*. Oxford: Oxford University Press.
Dressler, Wolfgang U.
1988 Language death. In Frederick J. Newmeyer (ed.), *Language: The socio-cultural context* (Language: The Cambridge survey, vol. IV), 184–192. Cambridge: Cambridge University Press.
Edwards, John
1992 Sociopolitical aspects of language maintenance and loss: A typology of minority language situations. In Willem Fase, Koen Jaspaert & Sjaak Kroon (eds.), *The maintenance and loss of minority languages*, 37–54. Amsterdam: Benjamins.
Farr, Marcia & Juyoung Song
2011 Language ideologies and policies: Multilingualism and education. *Language and Linguistics Compass* 5(9). 650–655.
Fishman, Joshua A.
1964 Language maintenance and language shift as a field of inquiry: A definition of the field and suggestions for its further development. *Linguistics: An Interdisciplinary Journal of the Language Sciences* 9. 32–70.
Fishman, Joshua A.
1965 Who speaks what language to whom and when. *La Linguistique* 2. 7–88.
Fishman, Joshua A. (ed.)
1966 *Language loyalty in the United States.* The Hague: Mouton and Co.
Fishman, Joshua A.
1991 *Reversing language shift.* Clevedon, UK: Multilingual Matters.
Fishman, Joshua A., Michael H. Gertner & Esther G. Lowy (eds.)
1985 *The rise and fall of the ethnic revival.* Berlin: Mouton.

Fishman, Joshua A., Robert L. Cooper & Andrew W. Conrad
1977 *The spread of English: The sociology of English as an additional language.* Rowley, MA: Newbury House.
Gal, Susan
1979 *Language shift: Social determinants of linguistic change in bilingual Austria.* San Francisco: Academic Press.
Gal, Susan
1996 Language shift. In Hans Goebl, Peter H. Nelde, Zdenek Stary & Wolfgang Wölck (eds.), *Contact linguistics*, vol. 1, 586–593. Berlin: Walter de Gruyter.
Garrett, Peter, Nikolas Coupland & Angie Williams
2003 *Investigating language attitudes: Social meanings of dialect, ethnicity, and performance*. Cardiff, UK: University of Wales Press.
Gibbons John & Elizabeth Ramirez
2004 Different beliefs: Beliefs and the maintenance of a minority language. *Journal of Language and Social Psychology* 23. 99–117.
Giles, Howard, Richard Y. Bourhis & Donald M. Taylor
1977 Towards a theory of language in ethnic group relations. In Howard Giles (ed.), *Language, ethnicity and intergroup relations*, 307–348. New York: Academic Press.
Gómez de García, Jule, Melissa Axelrod & Jordan Lachler
2009 English is the dead language: Native perspectives on bilingualism. In Paul V. Kroskrity & Margaret C. Field (eds.), *Native American language ideologies: Beliefs, practices and struggles in Indian Country*, 99–122. University of Arizona Press.
Grenoble, Lenore A. & Lindsay J. Whaley
1998 Toward a typology of language endangerment. In Lenore A. Grenoble & Lindsay J. Whaley (eds.), *Endangered languages*, 22–55. Cambridge: Cambridge University Press.
Grin, François, Tom Moring, Durk Gorter, Johan Häggman, Donald Ó Riagáin & Miquel Strubell
2002 *SMiLE: Support for language minorities in Europe*. Brussels: European Commission.
Gumperz, John J.
1982 *Discourse strategies*. Cambridge, UK: Cambridge University Press.
Haugen, Einar
1953 *The Norwegian language in America: A study in bilingual behavior*. University of Pennsylvania Press.
Haugen, Einar
1972 The ecology of language. In Anwar S. Dil (ed.), *The ecology of language: Essays by Einar Haugen*, 325–339. Stanford: Stanford University Press.
Heller, Monica (ed.)
2007 *Bilingualism: A social approach.* New York: Palgrave Macmillan.
Hickey, Raymond
2010 Contact and language shift. In Raymond Hickey (ed.), *The handbook of language contact*, 151–169. Wiley-Blackwell.
Hill, Jane H.
1978 Language death, language contact, and language evolution. In William C. McCormack & Stephen A. Wurm (eds.), *Approaches to language: Anthropological issues*, 45–78. The Hague: Mouton.
Hill, Jane H.
1983 Language death in Uto-Aztecan. *International Journal of American Linguistics* 49(3). 258–276.
Hill, Jane H.
2002 "Expert rhetorics" in advocacy for endangered languages: Who is listening, and what do they hear? *Journal of Linguistic Anthropology* 12(2). 119–133.

Himmelmann, Nikolaus P.
2010 Language endangerment scenarios: A case study from Northern Central Sulawesi. In Margaret Florey (ed.), *Endangered languages of Austronesia*, 45–72. New York: Oxford University Press.
Hoffman, Michol & James Walker
2010 Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change* 22. 37–67.
Kachru, Braj B.
1985 Standards, codification, and sociolinguistic realism: The English language in the outer circle. In Randolph Quirk & Henry G. Widdowson (eds.) *English in the world: Teaching and learning the language and literature.* Cambridge: Cambridge University Press.
Kandler, Anne, Roman Unger & James Steele
2010 Language shift, bilingualism and the future of Britain's Celtic languages. *Philosophical Transactions of the Royal Society B* 365. 3855–3864.
Kloss, Heinz
1966 German-American language maintenance efforts. In Joshua A. Fishman (ed.) *Language loyalty in the United States: The maintenance and perpetuation of non-English mother tongues by American ethnic and religious groups*, 206–252. The Hague: Mouton and Co.
Krauss, Michael
1992 The world's languages in crisis. *Language* 68. 4–10.
Kroskrity, Paul V. & Margaret C. Field
2009 *Native American language ideologies: Beliefs, practices and struggles in Indian Country.* Tucson: University of Arizona Press.
Kulick, Don
1992 *Language shift and cultural reproduction: Socialization, self, and syncretism in a Papua New Guinean village*. Cambridge: Cambridge University Press.
Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardner & Samuel Fillenbaum
1960 Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60. 44–51.
Le Page, Robert & Andrée Tabouret-Keller
1985 *Acts of identity.* Cambridge University Press.
Lewis, M. Paul & Gary F. Simons
2010 Assessing endangerment: Expanding Fishman's GIDS. *Romanian Review of Linguistics* 55(2). 103–20.
Lieberson, Stanley
1972 Bilingualism in Montreal: A demographic analysis. In Joshua A. Fishman (ed.), *Advances in sociology of language*, vol. 2, 231–54. The Hague: Mouton.
Lieberson, Stanley
1980 Procedures for improving sociolinguistic surveys of language maintenance and language shift. *International Journal of the Sociology of Language* 25. 11–27.
Li Wei
1994 *Three generations, two languages, one family.* Clevedon: Multilingual Matters.
Lüpke, Friederike
2015 Ideologies and typologies of language endangerment in Africa. In James Essegbey, Brent Henderson & Fiona McLaughlin (eds.), *Language documentation and endangerment in Africa*, 59–104. Amsterdam: John Benjamins.
Mackey, William F. & Donald G. Cartwright
1979 Geocoding language loss from census data. In William F. Mackey & Jacob Ornstein (eds.), *Sociolinguistic studies in language contact: Methods and cases*, 69–96. Berlin: De Gruyter.

McCarty, Teresa, Mary Eunice Romero-Little & Ofelia Zepeda
2006 Native language American youth discourses on language shift and retention: Ideological cross-currents and their implications for language planning. *International Journal of Bilingual Education and Bilingualism* 9(5). 659–677.
McLaughlin, Fiona
2008 The ascent of Wolof as an urban vernacular and national lingua franca in Senegal. In Cécile Vigouroux & Salikoko Mufwene (eds.), *Globalization and language vitality: Perspectives from Africa*, 142–170. London: Continuum.
Milroy, James & Lesley Milroy
1978 Belfast: Change and variation in an urban vernacular. In Peter Trudgill (ed.), *Sociolinguistic patterns in British English*, 19–36. London: Arnold.
Milroy, James & Lesley Milroy
2012 [1985] *Authority in language: Investigating Standard English.* New York: Routledge.
Musgrave, Simon & Michael C. Ewing
2006 Language and religion: A case study of two Ambonese communities. *International Journal of the Sociology of Language* 179. 179–94.
Nelde, Peter, Miquel Strubell & Glyn Williams
1996 *Euromosaic I: The production and reproduction of the minority language groups in the European Union*. Luxembourg: Office for Official Publications of the European Communities.
Odango, Emerzon Lopez
2015 *May Sasabihin ang Kabatan* 'The youth have something to say': Youth perspectives of language shift and linguistic identity. *Language Documentation and Conservation* 9. 32–58.
Ó Giollagáin, Conchúr, Seosamh Mac Donnacha, Fiona Ní Chualáin, AoifeNí Shéaghdha & Mary O'Brien
2007 Comprehensive linguistic study of the use of Irish in the Gaeltacht: Principal findings and recommendations. Dublin: Department of Community, Rural and Gaeltacht Affairs.
Palosaari, Naomi & Lyle Campbell
2011 Structural aspects of language endangerment. In Peter K. Austin & Julia Sallabank (eds.), *The Cambridge handbook of endangered languages*, 100–119. Cambridge, UK: Cambridge University Press.
Prince, Ellen
1988 On pragmatic change: The borrowing of discourse functions. *Journal of Pragmatics* 12. 505–18.
Sankoff, Gillian
2001 Linguistic outcomes of language contact. In Peter Trudgill, Jack Chambers & Natalie Schilling-Estes (eds.), *Handbook of sociolinguistics*, 638–68. Oxford: Basil Blackwell.
Sasse, Hans-Jürgen
1992 Language decay and contact-induced change: Similarities and differences. In Matthias Brenzinger (ed.), *Language death: Factual and theoretical explorations with special reference to East Africa*, 59–80. Berlin: Mouton de Gruyter.
Silva-Corvalán, Carmen
1994 *Language contact and change: Spanish in Los Angeles.* Oxford: Clarendon Press.
Smith-Hefner, Nancy
2009 Language shift, gender, and ideologies of modernity in Central Java, Indonesia. *Journal of Linguistic Anthropology* 19(1). 57–77.
Sorensen, Arthur
1967 Multilingualism in the northwest Amazon. *American Anthropologist* 69. 670–84.
Tollefson, James
1991 *Planning language, planning inequality.* New York: Longman.

Tsitsipis, Lukas D.
1989 Skewed performance and full performance in language obsolescence: The case of an Albanian variety. In Nancy Dorian (ed.), *Investigating obsolescence: Studies in language contraction and death*, 117–37. Cambridge: Cambridge University Press.
Vaish, Viniti
2007 Bilingualism without diglossia: The Indian community in Singapore. *International Journal of Bilingual Education and Bilingualism* 10(2). 171–187.
Vigouroux, Cécile & Salikoko Mufwene
2008 Colonization, globalization and language vitality in Africa: An introduction. In Cécile Vigouroux & Salikoko Mufwene (eds.), *Globalization and language vitality: Perspectives from Africa*, 1–31. London: Continuum.
Weinreich, Uriel
1953 *Languages in contact*. New York: Linguistic Circle of New York.
Winford, Donald
2003 *An introduction to contact linguistics.* Malden, MA: Wiley.
Zentella, Ana Celia
1997 *Growing up bilingual.* Oxford: Blackwell.

Maya Ravindranath Abtahian, Rochester (USA)

38. Language maintenance

1. Introduction
2. Studies of language maintenance
3. Maintaining ancestral languages
4. Conclusion
5. References

1. Introduction

Language maintenance "refers to a situation where a speech community, under circumstances that would seem to favour language shift, holds on to its language" (Coulmas 2013: 175). Most scholars use the terms 'language maintenance' and 'language shift' for opposing aspects of one and the same contact phenomenon (Appel and Muysken 1987; Brenzinger 1992, 2006; Fasold 1984; Fishman 1964; Pauwels 2016; Chapter 42 in this volume).

Studies on 'language shift' describe the processes and contexts of the decreasing use of a given language, as well as the receding competence in this language among a declining number of speakers, e.g. in Dorian's (1981) seminal study of Scottish Gaelic (see Chapter 37). Since then, compilations and global overviews of languages in danger have provided a better understanding of the extent of language shifts and of threats to global linguistic diversity (Brenzinger 2007a; ELCat 2013; Moseley 2007, 2010).

Hyltenstam and Stroud (1996: 567–568) distinguish between three basic types of studies on language maintenance and shift: (1) studies that focus upon the macro-societal framing conditions in language shift/maintenance; (2) studies that emphasize the connec-

https://doi.org/10.1515/9783110435351-038

tions between societal factors and the individual's actual perception of the situation as reflected in his/her language behavior; and (3) studies that focus on speakers' language competence and the structural consequences of the language maintenance/shift process. Much earlier sociolinguistic work tended to oversimplify and generalize language behavior, striving to establish bold categories and to design models which explained language choices and language use patterns by disregarding the multiplicity of motivations underlying individual language behavior. According to Smakman and Heinrich (2018: 2), "traditional approaches tend to ignore the social sphere of individuals and the many roles they play in it − as well as their sensitivity to the roles [… some of which] are subject to an individual's choice to abide by certain prescribed or proscribed roles and by that individual's communicative actions towards making their life a better one".

Furthermore, these approaches, as Smakman and Heinrich (2018: 2) point out, reduce motivations for language choices by individuals to a shared "collective characteristic while individualism in a globalising world is constantly making them more idiosyncratic and diverse and harder to capture as a group characteristic". Thus, while individuals decide for different reasons to abandon or to continue speaking their ancestral languages, the maintenance of these languages requires networks of speakers, namely language communities.

In discussing the maintenance of languages and the underlying motivations for such language choices, I distinguish between 'heritage languages', which are spoken by immigrants in the diaspora, and 'indigenous languages', which are non-dominant languages under threat in their home countries or regions. 'Ancestral languages' serves as a cover term for both kinds of community languages.

Threats to language maintenance vary and might derive, for example, from governmental policies. Such external pressures towards languages frequently trigger internal obstacles for the use of these languages. Speakers might anticipate revenues from linguistic assimilation and decide to abandon their languages. For speakers and communities, resisting linguistic assimilation often means accepting ongoing socioeconomic marginalization. Language choices in favor of language maintenance are based on the conscious decisions of individual speakers who decide to oppose pressures on the continued use of their ancestral languages. Many, often younger, community members, are no longer prepared to subscribe to a relentless struggle for the survival of ancestral languages. For example, the Jejueo revitalization movement, like many others, is mainly run and supported by elderly people (Brenzinger and Yang 2018).

The research on language maintenance must consider that power structures and the parameters of inequality have fundamentally changed in recent decades. As pointed out by Coulmas (2018: 22), "[s]ociolinguistics always has been and continues to be about inequality, but the relevant parameters of inequality nowadays are not the same as when urbanisation began in earnest. They have to do with new hierarchies of language and the influence of south-to-north migration and technology on life and communication in the city".

The claim made by Coulmas (2018: 12) that "[s]ociolinguistics is and has always been essentially urban" reflects the fact that more than half of the world's population now already lives in urban areas, and this will increase to two-thirds by 2050 (United Nations 2014). However, since most threatened indigenous languages are maintained by rural communities, most case studies of language maintenance of indigenous languages are from nonurban areas.

Language maintenance takes place in different contact settings. Immigrants might retain their heritage languages in the diaspora, or rural migrants might decide to continue using their ancestral tongues when moving to multilingual urban environments (cf. Chapter 27). Stable multilingualism is the main characteristic of successful language maintenance (cf. Chapter 22).

2. Studies of language maintenance

When Fishman (2013: 9) discusses his expanded understanding of sociolinguistics and the sociology of language as a discipline from 1964 onwards, he states that his "entire interest in topics related to this field began long before that and is strongly related to [his] commitment and activities pertaining to the maintenance of [his] own minority language − the Yiddish language and culture […]".

In 1951, Fishman and his family consciously decided to use Yiddish at home, with the explicit aim of maintaining their heritage language in their new US, English-speaking environment. Language maintenance, especially of endangered tongues, requires the kind of devotion to ancestral languages by individuals, families, and communities described by Fishman. Only the actual daily use of ancestral languages at home, with friends, and in the wider community ensures their survival. Language maintenance might seem to refer to a rather static language contact phenomenon, but it is precisely the opposite: maintaining non-dominant languages demands constant responses to ever-changing challenges from the speakers and language communities.

Fishman (1964) established the terms 'language maintenance' and 'language shift' in the 1960s. However, 40 years earlier, Kloss (1927) already used *Spracherhaltung* ('language maintenance') in his comparative studies of German in diaspora settings. Scholars such as Giles, Bourhis, and Taylor (1977) elaborated on models describing sociocultural factors that strengthen or weaken language maintenance. They identify 'demography', 'institutional support', and 'status' as major factors in defining 'ethnolinguistic vitality', a gauge of the strength of language maintenance among community members (see Clyne [2003] for a discussion of frameworks used to study language shift, maintenance and vitality). Coulmas (2013: 178−179), like other scholars, points out that "people act on the basis of perceptions, assumptions and assessments rather than facts". Language choices are therefore not predictable, since individuals make them on the basis of their subjective assessments, and not on objective grounds, as argued by Hyltenstam and Stroud (1996).

Fishman's (1991) *Graded Intergenerational Disruption Scale* (GIDS) and Lewis and Simons' (2009) *E(xpanded) GIDS* are tools designed for evaluating levels of language endangerment. UNESCO proposed the use of nine factors as key indicators for identifying different kinds of threats to language vitality (UNESCO 2003: 2). These factors attempt to capture the dynamics of changing patterns of language use. Factor (1), *intergenerational language transmission*, might be considered the most crucial indicator for language endangerment. When children no longer acquire an ancestral tongue, this language is obviously seriously threatened. Factor (2), *absolute number of speakers,* stresses the fact that small communities are vulnerable and incline towards language shift. Factor (3), the *proportion of speakers within a community*, helps to determine if an ancestral

language is still a core asset of a community's culture. Factor (4), *loss of domains,* and factor (5), *lack of response to new domains and media*, review changes in the use patterns of community languages. Factor (6), *availability of language teaching and learning tools,* also deals with aspects of the *Ausbau* of a language, such as the existence of a community orthography, literacy in the language, etc. Factor (7) addresses the *official stands of governments towards language diversity* and factor (8) evaluates the *speakers' attitudes towards their ancestral language*. Finally, factor (9) identifies the *amount and quality of language documentation* (Brenzinger 2007b: XI).

Applying these factors to analyzing an endangered language produces a detailed picture of the most severe threats to the maintenance of a specific language. This data in turn allows for the development of specific measures of language revitalization. Depending on the findings, the emphasis might be on language acquisition by children, development of language teaching materials, enhancement of the legal and/or perceived status of the language, or promotion of the use of the language in a greater number of domains.

3. Maintaining ancestral languages

The term 'language maintenance' might prompt associations of speakers and communities with being backward, purist, or conservative, or even with constituting a facet of nationalist ideologies. Most current language maintenance movements, however, are just the opposite. Speakers of indigenous and heritage languages claim linguistic rights and demand improved status of their ancestral languages as a means of empowerment, i.e. as measures to overcome discrimination and marginalization.

Safeguarding non-dominant languages requires individual speakers to actively resist linguistic assimilation. In order to ensure the future use of ancestral languages, meaningful roles for these languages in modern living contexts must be established, and the younger generation must be convinced to participate in community language maintenance and/or revitalization efforts.

While many indigenous languages (such as Breton in France, the Ryukyuan languages in Japan, Hawaiian in the US, and so on) were actively suppressed by governments in the past, some governments in more recent years have put measures aimed at maintaining and fostering multilingualism within national boundaries into place. The external support of language communities and individual speakers by provincial and national governments can play an important role in the maintenance of an ancestral language. 'Heritage language policies' are often endorsed to protect ancestral languages within nations. Most of these policies are directed at safeguarding indigenous languages, but some countries also make legal provisions for the maintenance of heritage languages of immigrants, as does, for example, the city-state of Singapore (Shah and Jain 2018).

Finally, global initiatives aiming to enhance multilingualism seem to be increasingly relevant, even for language choices made by rural communities in developing countries. International support, such as by UNESCO, is crucial, as maintaining a distinct linguistic identity often entails counteracting the aggressive behaviors and discriminatory policies of nation-states. However, most people still have no access to the internet, and in any case, the vast majority of the world's languages are not represented in this domain. The internet

will gain importance in the lives of all people, no matter where they live on our planet; UNESCO's activities directed toward establishing indigenous languages on the internet will therefore have a direct impact on the empowerment of their speakers (UNESCO 2017).

3.1. Heritage language maintenance in the diaspora

Heritage languages of immigrant groups are most often not endangered languages, as they remain vital or even dominant languages of states. Such immigrant languages can gain prestige in the diaspora, and proficiency in them might be rewarded by their countries of origin. More often, however, it can be observed that, while the first generation of immigrants maintains their heritage languages as a marker of identity (by recalling and claiming a shared common past), these sentimental motives tend to fade among the second generation, and more so among the third. Heritage languages seem to be chiefly maintained in close-knit communities and/or when they offer socioeconomic benefits to the individual speakers outside the community. The following case studies will illustrate differing language choices among immigrants, and the underlying motives.

I first discuss heritage language maintenance by contrasting German-speaking Namibian and South African communities, and then turn to language use among Gujarati immigrant communities in South Africa, East Africa, the UK (London), and Singapore. Without providing much historical background, I focus on describing the contexts, strategies, and motives for the language choices and use patterns that affect these heritage languages.

3.1.1. German in Namibia and South Africa

After more than a century, many descendants of German colonizers and settlers are still proficient in a distinct Namibian variety of German. The Namibian variety of German is the core marker of a distinct identity for the roughly 20,000 Namibian Germans who continue to speak their variety of German at home and in the community. Despite their small size, the German community contributes significantly to the Namibian economy and its cultural life.

Namibian-Germans have access to excellent education with German medium of instruction, to several German TV channels and Namibian-German radio broadcasts, and to the Namibian *Allgemeine Zeitung* ('General Newspaper'), the only daily newspaper printed in German on the African continent. Standard German is used in formal education, the media, and also in the tourist industry, as most foreign visitors to Namibia come from Germany (Shah and Zappen-Thomson 2018). At the same time, Namibian Germans often state that they feel alienated by Germans from Germany, and emphasize their African identity (Shah 2008). Standard German is therefore no marker of Namibian-German identity, and as a result, even the younger generation continues to speak the Namibian-German variety.

With far less access to such German-language infrastructure, German-speaking communities in South Africa have also retained distinct German varieties as community

languages. One of these communities resides in Kroondal, a small village in the North West province of South Africa. "As a small, rural German-speaking community, Kroondalers have by and large managed to resist language shift to Afrikaans and English, and have succeeded in maintaining their language for the past 130 years. The descendants of the earliest settlers are sixth generation German-speaking Kroondalers" (Shah, Herrmann, and Biberauer forthcoming). The Evangelical Lutheran church and the German primary school − both central institutions for the 400 Kroondalers − use predominantly German, which is also the main language at community gatherings. The maintenance of German language and identity aim at "[…] preserving their religious and social values based around a strong concept of family and being economically self-sufficient" (Shah, Herrmann, and Biberauer forthcoming).

In more recent years, however, the number of mixed marriages with mainly Afrikaans speakers has increased, and the hegemony of English in post-Apartheid South Africa does not stop at village boundaries. Both factors have put heavy pressure on the use of Kroondal German. Standard German remains popular, mainly because of the socioeconomic benefits which are offered by competence in this language. With intermarriage and increased mobility, there is little chance for future maintenance of the Kroondal German variety. Kroondal German wanes, along with the distinct German-South African Kroondal identity (Shah, Herrmann, and Biberauer forthcoming).

3.1.2. Gujarati in South Africa, East Africa, the UK, and Singapore

Many Indian immigrants to South Africa have retained their religious affiliations and maintained their cultural identities until today, but unlike the German immigrants mentioned above, they have abandoned their Indian heritage languages over the last 50 years.

Around the turn of the 19th century, after being on the African continent for almost half a century, Indian South Africans were still connecting to India and maintaining their Indian languages. For example, the *Indian Opinion*, a newspaper published by Mahatma Gandhi between 1903 and 1915, was written in Gujarati, Hindi, Tamil, and English. The newspaper is claimed to have inspired "anti-colonial, anti-imperial, anti-apartheid movements and movements in a quest for justice" (Dhupelia-Mesthrie 2011) not only for the Indian, but also for the black populations. Gandhi mentioned that the Gujarati sections were especially popular, and were read aloud in workplaces in Johannesburg (Dhupelia-Mesthrie 2011).

An account of the factors that led to the loss of these Indian languages in South Africa over the past 40 years is provided by Mesthrie (2007: 18):

> Shift took place because a change in ideology among young Indians bred solidarity and modernity across linguistic groups from the 1960s onwards. […] The silent linguistic revolution was perhaps part of a larger, quiet social revolution, which culminated in an acceptance of 'interlinguistic' marriages from the 1980s onwards and a new Indian South African identity. Indian South African English became the covert badge of identity representing the new solidarity. It signified that Indians were part of South Africa and not a transient Asian population identifying chiefly with India.

Different stages in the maintenance of Gujarati can be distinguished within the Gujarati community in the UK. Most of their ancestors had come from India to East Africa at

around the turn of the 19th century. From the 1960s onwards, when the East African British colonies gained independence, Gujaratis moved to the UK in large numbers.

As merchants and traders, Gujaratis set up small businesses in Kenya, Tanzania, and Uganda. They communicated in Swahili with their customers as well as with their African staff and domestic workers. In East Africa, Gujaratis were not exposed to 'Standard Gujarati', and used Swahili as the main language outside the community. This gave rise to an East African Gujarati (EA Gujarati) variety, characterized by heavy borrowing from Swahili. In Kenya, this EA Gujarati remains the main language of communication within the community, even among the younger generation (Shah 2013).

First and (albeit fewer) second generation London Gujaratis still maintain close family ties with their Kenyan relatives, with whom they communicate in EA Gujarati. Some elderly Indians in London remain monolingual in the EA Gujarati variety, and they enjoy a functioning spiritual, physical, and economic Gujarati-speaking infrastructure (Shah 2013).

First generation London Gujaratis had acquired Gujarati in East Africa, and the following generations learned EA Gujarati from them in the UK. These younger Gujaratis also received formal language instruction in Standard Gujarati, as spoken in India in community-run language schools. For these younger community members, there is less and less need to learn Gujarati, and little exposure through which to practice the language. Interest in formal Gujarati classes has been dramatically declining in recent years. Younger London Gujaratis associate themselves neither with East Africa nor with India, instead assuming a British Indian identity and communicating in English with each other. Thus, neither Standard nor EA Gujarati is being maintained (Shah 2016).

While the heritage languages of immigrant communities were distinct from their standard varieties in the previous examples, this seems not to be the case with Gujarati as spoken in Singapore, where Standard Gujarati is used among the Singapore Gujarati residents as well as the more recent immigrants from Gujarat.

Singapore is known for its bilingual language education policy, which demands that learners register for and learn an ethnically defined mother tongue in addition to English: "[w]hen Singapore gained independence in 1965, the government introduced four formative 'racial' categories, namely Chinese, Malay, Indian, and Others […]. [T]he city-state chose four official languages − Mandarin, Malay, Tamil, and English, with Malay in addition being designated the national language" (Shah and Jain 2018: 199). The choice of Tamil as the ethnic mother tongue for all immigrants from India was disputed, and Bengali, Gujarati, Hindi, Punjabi, and Urdu were added as additional language options. The Gujaratis are the smallest of these communities (4,124, census 2010), yet as merchants and traders, they have played a significant role in Singapore since their arrival at the end of the 19th century. They are also known for their efforts to maintain their cultural and linguistic heritage (Shah and Jain 2018: 202).

The Singapore Gujarati School offers language classes on the weekend, but even though the Gujarati population has increased in the city-state, the number of students enrolling in Gujarati classes has been consistently declining in recent years (Shah and Jain 2018: 206). Younger Gujaratis no longer speak Gujarati with their peers, and even less so with their families. While in the past, many Gujarati men married women from Gujarat in India, this is no longer common (Vinayak Dhansukhlal, personal communication [2016]). Intermarriage and the 'weakening of community bonds' have led to a rapid decline in the maintenance of Gujarati in Singapore. In particular, the more recent Guja-

rati immigrants, with uncertain residence status in Singapore, tend to choose Hindi as their mother tongue for the education of their children (Shah and Jain 2018: 214). Thus, even with supportive national language policies and community-run language teaching facilities, Gujarati will most likely not be maintained as a heritage language in Singapore.

3.2. Maintenance of indigenous languages

In contrast to heritage languages which usually continue to be spoken in their countries of origin, threatened indigenous languages become extinct when their speakers abandon them.

Evans (2010: 12) observes that indigenous languages of small communities have been maintained "wherever in the world societies have lain beyond the homogenizing reach of great empires". Evans (2010: 14) continuous: "[s]mall-scale societies in such parts of the world are economically self-sufficient, and proudly form the centre of their own social universe without needing to defer unduly to more powerful outside groups". The following case studies of the Akie in Tanzania and of the ǂKhomani San in the Northern Cape of South Africa contrast the maintenance of Akie with the revival of N|uu, the ancestral language of the ǂKhomani. The Akie are traditional hunter-gatherers who maintain a distinct linguistic and cultural identity. Despite their small numbers and low status in the region, and despite being subjugated by neighboring pastoralists and farmers for hundreds of years, they resist assimilation. The N|uu revival efforts by community members are attempts to revive and maintain their Indigenous Click language, which for them is a powerful testimony to their hunter-gatherer past.

3.2.1. Akie in Tanzania

In the 1920s, the British officer James Maguire declared that the language of the approximately 1000 Mosiro people (alias Akie; see König, Heine, and Legère 2017: 1) in the Maasai Steppe of Tanzania was on the verge of extinction (Maguire [1928] 1948: 10). Maguire describes the unstable bilingualism in which the 'Mosiro' increasingly abandoned their language in favor of the Maa language of the dominant neighbors, and states that the "language of the Mosiro is dying, as any language except Masai tends to do in the Masai country" (Maguire 1948: 10).

In defiance of the predicted demise of the language almost a century ago, more than 300 Akie are still fluent in this language today (König, Heine, and Legère 2017). One of the reasons for the maintenance of this language is the crucial role the language plays in communication with the ancestors (Heine, König, and Legère 2016: 329–330):

> […] [T]he life of all Akie […] is to a large extent determined by their relationship with the ancestors […], and the[se] *asííswe* have to be addressed only in the Akie language […]. Once one's father or mother dies s/he turns into an *asííswantee*, meaning that s/he is always watching you […]. In order not to be hit by accidents or death, people bless the ancestors at all special occasions, especially when something unexpected has happened or one intends to travel. Another complaint of the *asííswe* is that the Akie are rapidly discarding their traditional culture in favor of a *puuní* (Maasai) or *mɛ́ɛ́yɛ* ("Swahili") life style.

Heine, König, and Legère (2016: 331) foresee a stable quadrilingualism among the Akie, with Akie as the community's language, which is also used to communicate with the ancestors. Maasai is the language for engaging with the dominant pastoral neighbors. Nguu will gain importance, as − influenced by the Nguu − the Akie have recently started farming. Swahili is omnipresent in Tanzania, used in all official domains, such as education, media, and administration.

After they realized "that there is no place for them in the modern nation state of Tanzania" (Heine, König, and Legère 2016: 327), Akie seem "to have turned back to their own traditional world in which they feel at home" (Heine, König, and Legère 2016: 329). Maintenance of the Akie language and culture is seen by them as the only way which allows for their spiritual and physical survival (Heine, König, and Legère 2016: 327–329).

3.2.2. N|uu in the Northern Cape of South Africa

As of April 2019, N|uu − a Western variety of the N‖ng language cluster − is spoken by four elderly siblings living in different suburbs of Upington, a town in the Northern Cape province of South Africa. They consider themselves members of the ǂKhomani community. N|uu is the last remaining of the !Ui languages, which were once spoken by hunter-gatherer communities all over present-day South Africa (Güldemann and Vossen 2000). Afrikaans became the mother tongue of the ǂKhomani more than 70 years ago, when the white bosses on the farms forbid their workers from speaking their own languages.

Linguists had considered N‖ng extinct from the 1970s onwards (Traill 1999: 27), but in the late 1990s, Nigel Crawhall met a fluent speaker, Elsie Vaalbooi. Some twenty other elderly people in various parts of the Northern Cape revealed their competence in this language (Chamberlin and Namaseb 2001). Since then, Katrina Esau, one of the speakers, together with her granddaughter Claudia, have been teaching N|uu phrases and songs to children from their community and neighborhood.

The former hunter-gatherers and pastoral communities, as well as their languages, are often lumped together as 'Khoisan', ignoring the linguistic and cultural differences between them (Brenzinger 2014). What hunter-gatherers in fact share is a common past of economic exploitation, cultural marginalization, and even genocide. Assimilation of the hunter-gatherers into so-called 'colored' Afrikaans-speaking communities was one of the ways in which these languages and cultures, such as |Xam, disappeared. Genocide was another: "[w]herever they found themselves in their homeland, intruders either killed them or treated them with utter contempt and, through a process of 'taming', extinguished their |Xam identity" (Traill 1996: 183). In the second half of the 19th century, genocide committed by white Dutch-speaking pastoralists − joined by Baster, Griqua, and Khoi groups − wiped out entire hunter-gatherer communities (de Prada-Samper 2012: 186).

At the request of the last speakers, Sheena Shah and Matthias Brenzinger started to support N|uu language teaching efforts six years ago (Shah and Brenzinger 2017). In close collaboration with the community, and by considering the results from other lin-

guists, they helped to establish a N|uu community orthography. One of the largest phoneme inventories in the world, with 45 click phonemes, 30 non-click consonants, and 39 vowels, were represented on illustrated alphabet charts. They produced N|uu language posters, and published a trilingual reader in N|uu, Afrikaans, and English, which is now used in the community-run school (Shah and Brenzinger 2016).

Katrina Esau received the Order of the Baobab in silver by the president of South Africa in 2014 "[f]or her excellent contribution in the preservation of a language that is facing a threat of extinction. Her determination to make the project successful has inspired young generations to learn" (South African Government 2014). There is a growing interest among ǂKhomani to learn their ancestral language, but the lack of financial support and the fact that Katrina is the only teacher left make the reestablishment of N|uu as a medium of daily communication highly unlikely. N|uu might, however, be used symbolically instead to express a ǂKhomani San identity.

4. Conclusion

Individuals decide for various reasons whether or not to continue to use an ancestral language, but in either case, languages can only be maintained when being used within social networks of language communities. This requires cohesive communities with members who share a dedication to the maintenance of their ancestral languages. For use of the languages to be sustainable, young people in particular must choose to remain in (or join) such language communities. Communities defined by shared religion, for example, can encourage, or even demand, the use of ancestral languages from their members. Charismatic activists, such as Katrina Esau (mentioned above), as well as incentives, are essential to convincing people to engage in revitalization movements. Intermarriage, migration to other countries or from rural to urban areas, and increased mobility of community members are among the factors that weaken community bonds, and pose the most serious threats to language maintenance.

The case studies discussed above provide examples of the most common motivations for maintenance of ancestral languages. Sentimental, economic, and spiritual objectives prevail as motivations for retaining ancestral languages as core markers of individual and community identities. Reclaiming a lost hunter-gatherer ancestry is the main objective in the N|uu revival movement. Cultural and ritual practices, as well as family links to East Africa, seem to be the driving force for maintenance of EA Gujarati among older London Gujaratis. Kroondal German remains central to practicing a shared religious faith, at least for now. With access to mainstream society being denied, the Akie seem to retain their languages for lack of alternate opportunities. Maintenance of the Namibian-German variety is due to a strong local network within this immigrant community. New South African Indian and British Indian identities led to the abandonment of Indian ancestral languages in the second and third generations of immigrants, and in Singapore, Gujarati might soon be replaced by English, with Hindi becoming the ethnic language of education.

Heine, König, and Legère (2016: 313) rightly state that, while studies of the factors which lead to language replacement have received extensive scholarly attention, "[…] there is less information on why replacement does not take place − that is, why people

in certain situations do not give up their heritage language in favour of some other language even if their sociolinguistic environment discourages such a behaviour". The case studies discussed in this chapter provide insights into the complexities of language choices with regard to language maintenance.

5. References

Appel, Rene & Pieter Muysken
1987 *Language contact and bilingualism*. London: Edward Arnold.
Brenzinger, Matthias (ed.)
1992 *Language death: Factual and theoretical explorations with special reference to East Africa.* Berlin & New York: Mouton de Gruyter.
Brenzinger, Matthias
2006 Language maintenance and shift. In Keith Brown (ed.), *The encyclopedia of language and linguistics*, vol. 6, 542–548. Oxford: Elsevier.
Brenzinger, Matthias (ed.)
2007a *Language diversity endangered.* Berlin & New York: Mouton de Gruyter.
Brenzinger, Matthias
2007b Language endangerment throughout the world. In Matthias Brenzinger (ed.), *Language diversity endangered*, IX–XVII. Berlin & New York: Mouton de Gruyter.
Brenzinger, Matthias
2014 Classifying the non-Bantu click languages. In Lungisile Ntsebeza & Chris Sanders (eds.), *Papers from the pre-colonial Catalytic Project 1*, 80–102. Cape Town: University of Cape Town.
Brenzinger, Matthias & Changyong Yang
2018 Jejueo and language policies in South Korea. In Corinne Seals & Sheena Shah (eds.), *Heritage language policies around the world*, 185–198. London & New York: Routledge.
Chamberlin, J. Edward & Levi Namaseb
2001 Stories and songs across cultures: Perspective from Africa and the Americas. In Phyllis Franklin (ed.), *Modern Languages Association of America: Profession*, 24–38. New York: Modern Languages Association of America.
Clyne, Michael
2003 *Dynamics of language contact: English and immigrant languages*. Cambridge: Cambridge University Press.
Coulmas, Florian
2013 *Sociolinguistics: The study of speakers' choices*, 2nd edn. Cambridge: Cambridge University Press.
Coulmas, Florian
2018 Urbanisation and linguistic multitude. In Dick Smakman & Patrick Heinrich (eds.), *Urban sociolinguistics: The city as a linguistic process and experience*, 12–24. Abingdon, New York: Routledge.
de Prada-Samper, José Manuel
2012 The forgotten killing fields: "San" genocide and Louis Anthing's mission to Bushmanland, 1862–1863 (Review article of Mohamed Adhikari. 2010. *The Anatomy of a South African Genocide: The Extermination of the Cape San Peoples*). *Historia* 57(1). 172–187.
Dhupelia-Mesthrie, Uma
2011 The significance of Indian Opinion. *South African History Online*. http://www.sahistory.org.za/article/history-indian-opinion-newspaper/ (accessed on 18 December 2018).

Dorian, Nancy C.
1981 *Language Death: The life cycle of a Scottish Gaelic dialect*. Pennsylvania: University of Pennsylvania Press.
ELCat
2013 *the Endangered Language Project.* http://www.endangeredlanguages.com/ (accessed on 18 December 2018)
Evans, Nicholas
2010 *Dying words: Endangered languages and what they have to tell us*. Chichester: Wiley-Blackwell.
Fasold, Ralph
1984 *The sociolinguistics of society.* Oxford: Basil Blackwell.
Fishman, Joshua A.
1964 Language maintenance and language shift as a field of inquiry: A definition of the field and suggestions for its further development. *Linguistics* 2(9). 32–70.
Fishman, Joshua A.
1991 *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon-England: Multilingual Matters.
Fishman, Joshua A.
2013 Introductory comments by the author. *Linguistics* 51. 9–10.
Giles, Howard, Richard Y. Bourhis & Donald M. Taylor
1977 Towards a theory of language in ethnic group relations. In Howard Giles (ed.), *Language, ethnicity and intergroup relations*, 307–348. London: Academic Press.
Güldemann, Tom & Rainer Vossen
2000 Khoisan. In Bernd Heine & Derek Nurse (eds.), *African languages*: *An introduction*, 99–122. Cambridge: Cambridge University Press.
Heine, Bernd, Christa König & Karsten Legère
2016 Reacting to language endangerment: The Akie of north-central Tanzania. In Martin Pütz & Luna Filipović (eds.), *Endangered languages: Issues of ecology, policy and documentation* (IMPACT: Studies in Language and Society 42), 313–333. Amsterdam, Philadelphia: John Benjamins.
Hyltenstam, Kenneth & Christopher Stroud
1996 Language maintenance. In Hans Goebl, Peter H. Nelde, Zdenek Stary & Wolfgang Wölck (eds.), *Kontaktlinguistik. Contact linguistics*. *Linguistique de contact*, vol. 1., 567–578. Berlin: Walter de Gruyter.
Kloss, Heinz
1927 Spracherhaltung. *Archiv für Politik und Geschichte* 8. 456–462.
König, Christa, Bernd Heine & Karsten Legère
2017 Akie as a Language Island. Conference paper presented at Language Islands in Bad Münster, Germany June 14–18, 2017.
Lewis, Paul M. & Gary F. Simons
2009 Assessing endangerment: Expanding Fishman's gids. *Revue Roumaine de Linguistique* 55. 1–30.
Maguire, R. A. J.
1948 [1928] Il-Torobo. *Tanganyika Notes and Records* 25. 1–27.
Mesthrie, Rajend
2007 Language shift, cultural change and identity retention: Indian South Africans in the 1960s and beyond. *South African Historical Journal* 57. 134–152.
Moseley, Christopher (ed.)
2007 *Encyclopedia of the world's endangered languages*. New York: Routledge.
Moseley, Christopher (ed.)
2010 *Atlas of the world's languages in danger*, 3rd edn. Paris: UNESCO Publishing.

Pauwels, Anne
2016 *Language maintenance and shift.* Cambridge: Cambridge University Press.
Shah, Sheena
2008 Negotiating multilingual identities in interaction: The Südwesters of Namibia. In Otilia Arnandiz & Maria Safont Jorda (eds.), *Achieving multilingualism: Wills and ways*. *Proceedings of the 1st International Conference on Multilingualism*, 389–402. Castelló de la Plana: Publicacions Universitat Jaume I.
Shah, Sheena
2013 *Factors affecting proficiency among Gujarati heritage language learners on three continents.* Washington, DC: Georgetown University dissertation.
Shah, Sheena
2016 Negotiating Gujarati identity through language in Kenya and the UK. Paper presented at the 21st Sociolinguistics Symposium, University of Murcia, Spain, 15–18 June.
Shah, Sheena & Marianne Zappen-Thomson
2018 German in Namibia. In Corinne A. Seals & Sheena Shah (eds.), *Heritage language policies around the world*, 128–147. Abingdon & New York: Routledge.
Shah, Sheena & Matthias Brenzinger
2016 *Ouma Geelmeid ke kx'u ǁxaǁxa N!uu ◊ Ouma Geelmeid gee N!uu ◊ Ouma Geelmeid teaches N!uu.* Cape Town: CALDi. https://open.uct.ac.za/handle/11427/17432/ (accessed on 18 December 2018).
Shah, Sheena & Matthias Brenzinger
2017 Writing for speaking: The N!uu orthography. In Mari Jones & Damien Mooney (eds.), *Creating orthographies for endangered languages*, 109–125. Cambridge: Cambridge University Press.
Shah, Sheena & Ritu Jain
2018 Gujarati in Singapore. In Corinne A. Seals & Sheena Shah (eds.), *Heritage language policies around the world*, 199–217. Abingdon & New York: Routledge.
Shah, Sheena, Erika Herrmann & Theresa Biberauer
Forthcoming Kroondal Deutsch. In Hans Boas, Ana Deumert, Mark L. Louden & Péter Maitz (eds.), *Varieties of German worldwide*. Oxford: Oxford University Press.
Smakman, Dick & Patrick Heinrich
2018 Introduction. Why cities matter for a globalizing sociolinguistics. In Dick Smakman & Patrick Heinrich (eds.), *Urban sociolinguistics: The city as a linguistic process and experience*, 1–11. Abingdon, New York: Routledge.
The National Orders
April 2014 www.thepresidency.gov.za/file/1013/download?token=IT98_iDg (accessed on 12 September 2018).
Traill, Anthony
1996 !Khwa-Ka Hhouiten Hhouiten "The Rush of the Storm": The linguistic death of /Xam. In Pippa Skotnes (ed.), *Miscast: Negotiating the presence of the Bushmen*, 161–183. Cape Town: University of Cape Town Press.
Traill, Anthony
1999 *Extinct South African Khoisan languages* [CD and booklet]. Johannesburg: University of the Witwatersrand.
UNESCO
2003 Language vitality and endangerment. https://ich.unesco.org/doc/src/00120-EN.pdf (accessed on 12 September 2018).
UNESCO
2017 Linguistic diversity and multilingualism on Internet http://www.unesco.org/new/en/communication-and-information/access-to-knowledge/linguistic-diversity-and-multilingualism-on-internet/ (accessed on 12 September 2018).

United Nations
2014 World's population. http://www.un.org/en/development/desa/news/population/world-urbanization-prospects-2014.html/ (accessed on 12 September 2018).

Matthias Brenzinger, Bloemfontein (South Africa)

IV. Methodological issues

39. Quantitative analysis of language contact data

1. Introduction
2. Preliminary concepts
3. Descriptive statistics
4. Correspondence analysis
5. Other descriptive multivariate techniques
6. Hypothesis testing
7. Software programs
8. References

1. Introduction

The goal of quantitative analysis is to uncover the structure that underlies a set of observations. These observations are often quantitative measurements, such as vowel durations or Likert scale measures; they can also be qualitative characteristics, such as information about speakers' language choices or their choice of a lexical form. Whichever type of observation is being considered, and whatever the size and complexity of the database, skill and care are needed to ensure that the analysis is carried out with the appropriate objective procedures. In a broad sense, data analysis encompasses a range of activities, starting with the conception of a study and ending with the presentation of the final results. This chapter focuses on the exploration and analysis steps in this undertaking and, in particular, on those strategies and statistical methods appropriate for the exploration and analysis of language contact data.

Linguistics has paid growing attention to the role of quantitative analysis. A number of books about statistics have been written specifically for linguists. Woods, Fletcher, and Hughes (1986) summarize fundamental statistical concepts useful for students. Rasinger (2008) is a readable introduction to quantitative analysis, with discussion of research design and questionnaire construction. More detailed presentations include Eddington (2015), Johnson (2008), Larson-Hall (2016), and Rietveld and van Hout (1993). In addition, certain domains of linguistic inquiry, such as experimental psycholinguistics, variationist sociolinguistics, and second language acquisition, have devoted considerable energy to establishing best practices in quantitative analysis, and these guidelines serve as a useful source of methodological information.

Research on language contact has followed different approaches, and researchers have applied a fairly wide variety of statistical techniques. Many tend to use simple descriptive statistical tools, such as a frequency table or a display of means. Nevertheless, in earlier studies, a small handful of researchers adopted advanced statistical methods, such as factor analysis (Ma and Herasimchuk 1971) and cluster analysis (Le Page et al. 1974). Subsequent contributions include other multivariate techniques, such as variable rule analysis and correspondence analysis. More recently, mixed-effects models, random forests, and conditional inference trees have been added to the toolkit.

This growth in the variety and sophistication of statistical methods used for the quantitative analysis of language contact data parallels increasing interest in different sources of data. Studies are based on the analysis not only of spontaneous oral interviews and large databases of oral production, but also of text messages, chat conversations, newspa-

https://doi.org/10.1515/9783110435351-039

pers, and other forms of written correspondence. Some approaches focus on data from experimental research and Natural Language Processing. Growth in the use of statistical methods also parallels the development of software, often open-source, which is designed for the exploration and analysis of large linguistic corpora. Furthermore, recent advances in statistics have introduced new analytical techniques that can deal with issues − such as sparse data or high covariance among predictor variables − that often arise in the study of linguistic data and can be difficult to analyze using classical statistical methods.

This chapter gives a practical overview of the quantitative methods found in the language contact literature. The discussion surveys different statistical techniques that have proven to be useful; it is, by necessity, not exhaustive. The overview includes the analysis of data from small-scale and large-scale studies that are based on linguistic interviews, language-use questionnaires, and large databases. It does not examine the quantification of language census data or other national surveys. The presentation includes methods of elementary univariate description, more advanced multivariate techniques, and hypothesis testing. Detailed applications of the chi-square test and correspondence analysis are presented for illustrative purposes.

2. Preliminary concepts

2.1. Types of data

Measurement theory recognizes four data types: categorical and ordinal (which are generally referred to as qualitative data), and interval and ratio data (which are called quantitative data).

Categorical data are labels of attributes that assign an observation to a particular category or group. Examples are gender (male, female, other), ethnicity (Caucasian, Asian, African), and language choice (English, French, German). Much of the linguistic data studied by linguists is categorical. Some examples are phonological context (syllable initial versus syllable final), alternate simple past tense forms of *see* (*saw*, *seen*), and lexical variants (*automobile*, *car*, *vehicle*). These categories or attributes can be counted, but they cannot be ordered nor can they be used in mathematical formulas.

In ordinal measurements, responses to a variable receive numbers that order them with respect to some criterion. Replies to the statement 'You speak Gaelic to your parents' − (1) never, (2) sometimes, (3) often, (4) usually, (5) always − are ordinal data. There is an ordered relation among the possible replies, but the intervals between levels are not equal. The numbers can be used in very specific mathematical formulas.

Interval and ratio data have units that are not only ordered, but also equidistant from each other. A semantic differential with a five-point scale constitutes an interval measure. Phonetic measurements, such as voice onset time (VOT) durations (measured in milliseconds) or vowel formant frequencies (measured in Hertz), are ratio measurements. These two data types differ in that ratio data have an origin, or zero, whereas interval data do not. Many researchers refer to interval and ratio data as continuous data because both data types are analyzed with the same statistical techniques.

It is important to distinguish among the four data types because different statistical procedures are appropriate for the analysis of each type. Indeed, the choice of a statistical procedure can become complicated in cases where a study has several data types; for example, the observed data may be quantitative type data, while the predicting factors may be both qualitative and quantitative types. Many of the variables studied in language contact research are qualitative (either categorical or ordinal), and this limits the choice of data analytical methods that can be applied.

2.2. Matrix representation

The observations pertinent to data analysis are arranged in the form of a data matrix, which looks like a spreadsheet. The arrangement of the matrix is described in terms of rows and columns, which helps to bring order to the data that are under study.

A typical arrangement is the subjects-by-variables matrix shown in Table 39.1. The rows are individual subjects or participants in a study, and the columns are variables or characteristics that provide information about each subject. These variables can be social or ethnographic details, such as participant identification number, gender, ethnicity, age, and/or linguistic information, such as replies to a language choice questionnaire or use of a particular phonetic, lexical, or syntactic feature. Clearly, the variables can be of any of the four data types. The intersection of a row and a column is the information given by a specific variable for a specific individual.

The columns in the matrix are generally referred to as variables. Many research projects distinguish between independent and dependent variables. In these studies, the main research question might be stated as: how do one or more of the independent variables affect the dependent variable? Dependent variables are the variables that the researcher wishes to explain; independent variables are those features that can influence a dependent variable. In other research projects, however, interest may be in finding similarities and differences among a number of variables. In this case no dependent versus independent distinction is made.

Other setups of the data matrix are possible. In a tokens-by-variables matrix, the rows are linguistic tokens − individual occurrences of a particular feature − and the columns describe the social and linguistic contexts in which each token occurs. It is also possible

Tab. 39.1: Model of a subjects-by-variables data matrix of dimension (n × 5). Each entry v_{ik} represents the value of variable Vk for subject i

		columns: the variables				
		V1	V2	V3	V4	V5
rows: the subjects	1	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}
	2	v_{21}	v_{22}	v_{23}	v_{24}	v_{25}
	…			…		
	i	v_{i1}	v_{i2}	v_{i3}	v_{i4}	v_{i5}
	…			…		
	n	v_{n1}	v_{n2}	v_{n3}	v_{n4}	v_{n5}

to aggregate information about individual members of social groups to formulate a groups-by-variables matrix. In any of its possible forms, the data matrix serves both as an input to statistical analysis, and as a presentation of the data under study.

3. Descriptive statistics

The analysis of the data matrix begins with a description of the individual variables in the matrix. This is the case for univariate studies, where one variable is of interest, as well as for multivariate studies, where many variables are studied simultaneously. Qualitative and quantitative data should be distinguished in both types of study. Increasingly, graphics are used to communicate this descriptive statistical information. See Larson-Hall (2017) for a discussion of data visualization and of graphic alternatives to the three graphics discussed in this section.

3.1. Description of a single variable

The most common descriptor of categorical type data is a frequency distribution. This distribution records the relative proportion of the population that belongs to each category of a variable. This information is generally reported in a one-way frequency table or by graphical techniques, such as a bar chart. In their study of English borrowings in French spoken in Ottawa-Hull, Canada, Poplack, Sankoff, and Miller (1988) present frequency information about English borrowings by giving a list of percentages by grammatical category. The bar chart in Figure 39.1 is a graphical presentation of these 'borrowability' data.

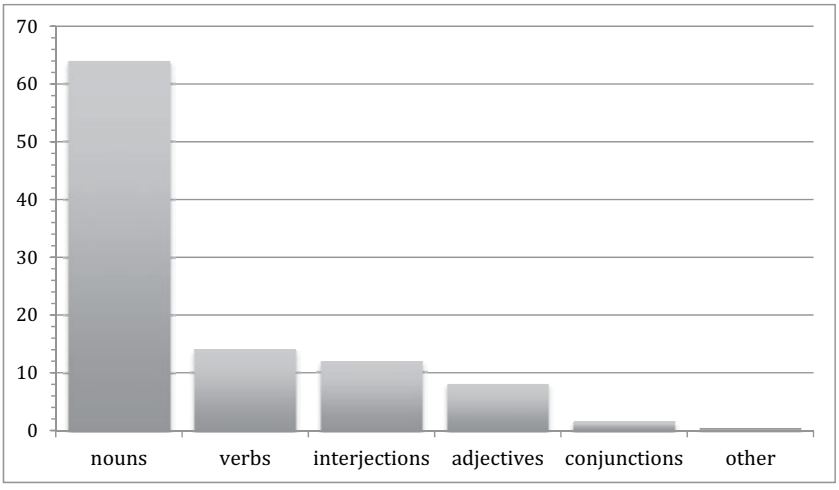


Fig. 39.1: Bar chart showing the proportion (in %) of English borrowings by grammatical category in French spoken in Ottawa-Hull (Poplack, Sankoff, and Miller 1988)

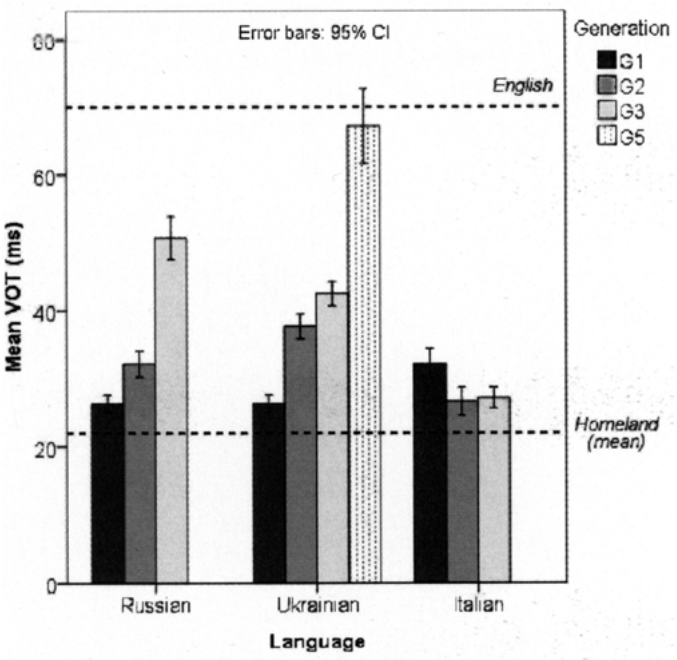


Fig. 39.2: Barplots showing means and 95 % confidence intervals of VOTs (in msec) for several generations of three heritage languages spoken in Toronto. Mean English VOT value is also plotted for comparison (Hrycyna et al. 2011: 166)

Three descriptive statistics are used to summarize information about ordered, interval, and ratio type data. These are (1) mean, median, and mode, which are measures of central tendency; (2) standard deviation, which is a measure of variability or dispersion; and (3) the shape of the distribution, which can have features such as symmetrical, skewed, and bimodal. While the third descriptor is only reported infrequently, the first two descriptive statistics − measures of central tendency and variability − are almost always given. Indeed, standard deviation and variance (which is calculated as the square of the standard deviation) measure how spread out the data are, and the theme of identifying the sources of this spread or dispersion in the data is central to quantitative analysis. The barplot and the boxplot are two common graphical presentations of descriptive statistics that characterize non-categorical data.

The barplots in Figure 39.2 give descriptive statistical information about the voice onset time (VOT) of stop consonants in three heritage languages − Russian, Ukrainian, and Italian − spoken in Toronto, Canada (Hrycyna et al. 2011). Each vertical bar in the figure provides a measure of the central tendency, or mean, of the VOTs of a specific group of speakers. The error bars around each mean give the 95 % confidence interval; this interval is the range of values in which lies the true mean, assuming that the data follow a symmetrical (normal) distribution. While barplots do not show information about the dispersion or shape of the distribution of the data, they are useful when the means of several variables − such as language and generation − need to be presented on one graph.

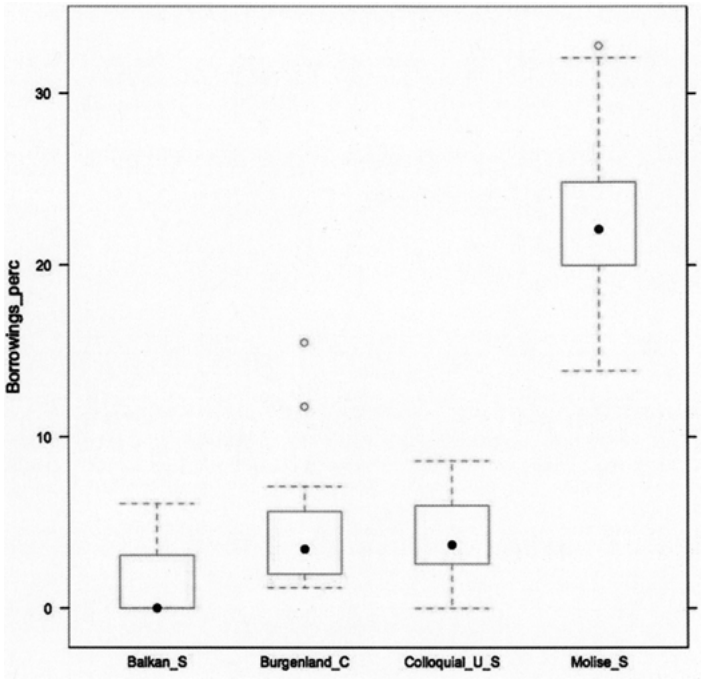


Fig. 39.3: Boxplots showing rates of noun borrowing among minority Slavic language speakers in four languages: Balkan Slavic, Burgenland Croatian, Colloquial Upper Sorbian, and Molise Slavic (Adamou et al. 2016: 532)

A different, and more informative, presentation of descriptive statistical information is the boxplot, also called a box-and-whisker diagram. Figure 39.3 is a graph of noun-borrowing rates among Slavic minority speakers who reside in Austria, Germany, Greece, and Italy (Adamou et al. 2016). The box for each group of speakers presents the median and the interquartile range (50 % of the data points); the whiskers extend to 1.5 times the interquartile range above the upper and below the lower quartiles; circles represent individual outliers. Visual inspection shows that the Molise Slavic speakers (in Italy) have a 'high' average borrowing rate compared with speakers from the three other languages, who have 'low' average borrowing rates; furthermore, Molise Slavic borrowing rates have the greatest degree of variability. Among the three 'low borrower' languages, only the Burgenland Croatian speakers (in Austria) have some outliers. Finally, the Balkan Slavic language (in Greece) has a lopsided spread in borrowing rates; this skewed distribution, which is different from the relatively symmetrical distributions observed in the other languages, is due to the large number of speakers with no noun borrowings. In sum, boxplots provide a compact visualization of summary information about central tendency and dispersion in a dataset, and about the shape of the data distribution.

3.2. Associations between two variables

The association between two qualitative variables is often represented by contingency and frequency tables. Contingency tables are cross-tabulations that display the number of times that each combination of two variables occurs; one variable is in rows and the other in columns. A common application of contingency and frequency tables is in questionnaires that document language choices. For example, different situations of language use are usually cross-tabulated with relative frequency of use (always, sometimes, never). See Dorian (1981: 179–187) for examples of questionnaires about Gaelic-English contact in Scotland.

Associations between two quantitative variables can be displayed graphically with scatterplots. This association can be measured by a linear correlation coefficient, such as Pearson's *r*, which informs about the nature and direction of the relationship between the variables (direct or indirect) and about the strength of the relationship. However, interpretation of the *r* statistic requires reference to the scatter diagram in order to recognize whether the relationship is a straight line, a curve, or a less well-organized cloud of points. Cichocki (2008) correlates relative frequency of occurrence of an English-type (retroflex) /r/ consonant in French-origin words with the concentration of francophone speakers in 18 Acadian French localities in eastern Canada. Visual inspection of the scatterplot indicates a curvilinear relationship between these variables.

The association between a qualitative and a quantitative variable can be described by examining the quantitative variable at the different levels of the categorical variable. The barplots in Figure 39.2 illustrate VOT durations (a quantitative variable) for three generations (levels) of Russian speakers, for four generations of Ukrainian speakers, and for three generations of Italian speakers (all qualitative variables).

3.3. Chi-square and the association between two categorical variables

Because language contact data are often qualitative, we discuss the association between two categorical variables in more depth. Table 39.2 is a (2 × 3) matrix that documents language choices made by a group of respondents in their communication with their spouses. The data are from a survey carried out in 2000 by Choi (2005) to replicate Rubin's well-known 1960 study of Guarani-Spanish language contact in Paraguay (Rubin 1968). The rows are two points in time: 1960 and 2000; the columns are three language choices: Spanish, Guarani, and both Spanish and Guarani. The entries are the number of respondents, at each of the two points in time, who indicated a particular language choice in communicating with their spouse.

Tab. 39.2: Frequency-table data for language choices in communication with one's spouse (Choi 2005: 239)

Year	Spanish	Guarani	Both	*Total*
1960	10 (18.9 %)	16 (30.2 %)	27 (50.9 %)	*53*
2000	20 (32.8 %)	6 (9.8 %)	35 (57.4 %)	*61*

To interpret these data, Choi calculated the row percentages, which have been added to the table of frequencies. She observes three differences in language choices between 1960 and 2000: greater use of Spanish (an increase from 18.9 % to 32.8 %), a drop in use of Guarani (30.2 % to 9.8 %), and a slight increase in bilingual use (50.9 % to 57.4 %).

We analyze this small dataset using the chi-square (χ^2) statistic, which is the most widely used measure of association between the rows and columns of a contingency table. This statistic tests the null hypothesis, i.e. that there is no association between rows and columns. The formula used to calculate χ^2 compares the expected and observed frequencies in each cell; cells with large differences between these two frequencies contribute a greater amount to the statistic than do cells with smaller differences between these two frequencies. In Table 39.2 the value of χ^2 is 15.13, and when compared with reference values given in standard tables, this number turns out to be significant; in fact, the level of significance is very high. This result is reported as follows: χ^2 (2 degrees of freedom, N = 113) = 15.13, $p < 0.0001$. This means that the observed data do not support the null hypothesis, and we can conclude that there is a statistically significant relationship between the two points in time and language choice in Paraguay.

The analysis of frequency-contingency tables has some noteworthy drawbacks. One is that the chi-square statistic does not identify the pattern or patterns that make up the relations between the two categorical variables. A systematic cell-by-cell examination of differences between expected and observed frequencies − and more specifically, an inspection of Pearson residuals and adjusted residuals − can go a long way to helping the researcher identify major associations between rows and columns. However, analyzing large tables, with say 10 rows and 10 columns, becomes cumbersome and complicated. Such tables are hard to display, and their interpretation may not be clear. Other techniques, such as correspondence analysis (see below), can be helpful in this respect.

Finally, while the chi-square statistic reveals whether or not there is an association between the two variables, it does not provide information about the strength of that association. A useful index of effect sizes is Cramer's V; this statistic is based on the chi-square statistic and gives a value between 0 and 1. In general, values near 1 indicate a large effect, while those near 0 indicate a very small effect. Bullock, Serigos, and Toribio (2016: 186) use Cramer's V to compare the proportion of English borrowings (nouns and adjectives) to Spanish nouns and adjectives in Spanish-language newspapers which are aimed at readers from different socioeconomic groups in Argentina and in Puerto Rico. The value of Cramer's V is 0.068 for the Argentine data, and 0.032 for the Puerto Rican data, indicating that both associations are very weak. However, there is a greater difference in borrowing counts among the Argentine newspapers than there is among the Puerto Rican newspapers, and this leads the authors to speculate that English borrowings in Argentina may be more closely related to social class than in Puerto Rico.

4. Correspondence analysis

Correspondence analysis is a descriptive and exploratory technique that analyzes qualitative type data. It is also known as dual scaling, Hayashi's quantification, homogeneity analysis, and optimal scaling (Greenacre 2017). The basic strategy is to represent the information found in a data matrix in the form of a graph that can be used for data

exploration. The technique makes no assumptions about any mathematical hypotheses on the structure of the data, nor does it require that the data sets be large. Several descriptive statistical indices accompany the outputs, and these help the researcher to find patterns and trends in the data and to identify individual outliers or deviant data entries.

There have been two main types of application of this technique to the exploratory study of language contact data: multiple correspondence analysis (in which the variables are categorical type data), and the analysis of frequency or contingency tables (which are cross-tabulations of two or more categorical variables). Because categorical data occur frequently in language contact studies, we present an illustrative example of the second type of application. See Cichocki (1986, 1996) for a detailed description of correspondence analysis (specifically, dual scaling) and several applications to the analysis of linguistic variation and language contact data.

4.1. Multiple correspondence analysis

This analysis is sometimes called multiple-choice data. In a typical application, subjects reply to questions that have multiple-choice answers, and the data matrix has a subjects-by-answers structure.

Inoue (1993) analyzes the prosodic integration of German and English borrowings into Japanese. The variables in his study include the borrowing, its pitch accent, and socio-psychological attributes of the speakers. The research describes variation in the type of pitch accent assigned to a borrowing: a flat (unmarked) pitch accent is used by speakers who have a high familiarity with the semantic field from which the word comes; a high (marked) pitch accent occurs for borrowings that belong to a semantic field with which speakers are less familiar.

4.2. Correspondence analysis of a frequency table

The data matrix in Table 39.3 is based on Choi's (2005) study of Guarani-Spanish language choices in Paraguay. The entries in the matrix are frequencies of language choices in five domains (communication with spouse, parents, children, neighbors, and a priest) at two points in time (1960 and 2000). There are three language choices: Guarani, Spanish, and both Guarani and Spanish. The data form a (10 × 3) contingency-frequency matrix.

Based on the chi-square statistic calculated for this table, the hypothesis that the rows and columns are statistically independent in this population should be rejected. According to the chi-square test: χ^2 (18 degrees of freedom, N = 629) = 203.46, $p < .0001$. Hence, the inference can be made that language choices, domains, and points in time are associated. We apply correspondence analysis to this data matrix to explore the makeup of this association.

Analysis yields the scatterplot shown in Figure 39.4. This diagram is called a biplot, which is the overlay of the plane formed by the rows (i.e. the domains/years) and the plane formed by the columns (i.e. the three language choices). The plane of the three column points resembles a triangle that provides a frame of reference for interpreting

Tab. 39.3: Frequency-table data for language choices in five domains at two points in time (Choi 2005: 239)

domain/year	Spanish	Guarani	Both	*total*
spouse 1960	10	16	27	*53*
spouse 2000	20	6	35	*61*
parents 1960	10	37	18	*65*
parents 2000	13	28	29	*70*
kids 1960	34	6	18	*58*
kids 2000	31	2	28	*61*
neighbors 1960	11	32	20	*63*
neighbors 2000	12	13	46	*71*
priest 1960	29	12	16	*57*
priest 2000	55	3	12	*70*
total	*225*	*155*	*249*	*629*

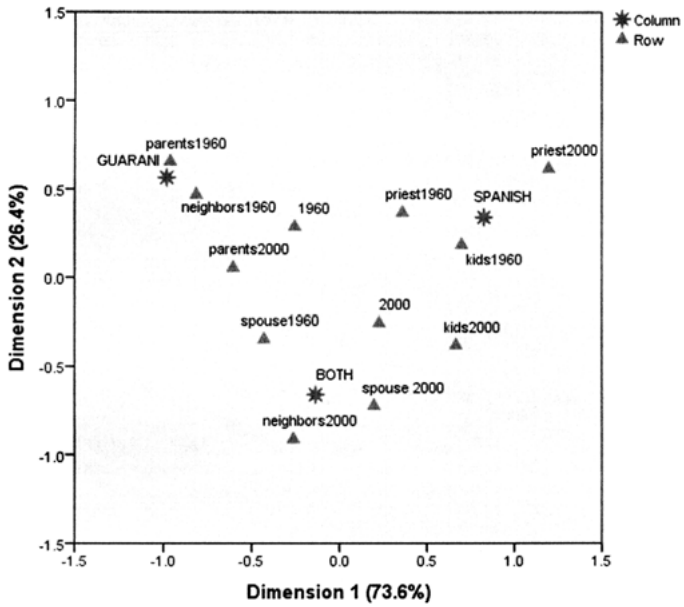


Fig. 39.4: Graphical display of the correspondence analysis of the (10 × 3) frequency table in Table 39.3. This display superposes the domains/years (rows in the matrix) space and the language choices (columns) space

the row points. The ten domains/years form the row plane; proximity or distance between two row points indicates the degree of similarity or dissimilarity between those rows, respectively.

In looking at the biplot, we can make a number of observations. For example, in 1960, language choices with parents and neighbors were very similar and were close to Guarani. In 2000, language choices in these domains were different, with both domains

showing movement in the direction of 'both' languages. The shift in communication with neighbors between 1960 and 2000 is noticeably large. Language choices in communication with a priest were close to Spanish in both 1960 and 2000, and they moved even further away from Guarani in 2000. Language choices with children moved away from Spanish to a choice of both languages.

The two-dimensional plane in Figure 39.4 accounts for all of the variance in the data matrix (referred to as 'inertia' in correspondence analysis). Descriptive statistics − provided by correspondence analysis − measure the variance for each dimension, and indicate that the horizontal axis accounts for a larger proportion of variance than does the vertical axis: 73.6 % versus 26.4 %. This shows that the shift from Guarani to Spanish is the major pattern in these data.

Other descriptive statistics point to cells in the data matrix that contribute heavily to the significant chi-square statistic calculated for the contingency-frequency table. These include key elements that underlie the associations between domains/years and language choices. Three points play a dominant role in determining the horizontal distribution of points (Dimension 1) that contrasts Guarani and Spanish language choices: 'parents' and 'neighbors' in 1960, at the Guarani end, and 'priest' in 2000, at the Spanish end. Points that determine the vertical distribution (Dimension 2) are 'neighbors' and 'spouse' in 2000, at the lower or 'both Guarani and Spanish' end of the graph, and 'parents 1960' and 'priest 2000', at the end opposite the 'both' language choice.

The analysis reveals general patterns about the two variables domains and years. The domains with the largest displacements on the biplot (i.e. that show the greatest degree of movement in language choice) are 'neighbors', followed by 'priest' and then 'parents'. The two points labeled '1960' and '2000' in Figure 39.4 summarize the overall movement in language choices between the two years. These points, called 'supplementary points' in the terminology of correspondence analysis, represent the centroids of language choices for the five domains at each of the two points in time. According to the positions of these points in the figure, language choices across time are moving away from Guarani toward greater use of Spanish (the main movement), but also toward greater bilingual use (a secondary movement). Clearly, language choices made over time do not follow a linear or one-dimensional path away from Guarani.

5. Other descriptive multivariate techniques

A number of other multivariate techniques have been used to identify patterns that underlie a set of variables. These analyses aim to explore relations or associations among the variables and not to analyze them in the context of a dependent-independent relation. The variables can be qualitative or quantitative. Two related techniques, not discussed here, that are used to analyze quantitative variables are factor analysis and principal components analysis.

5.1. Implicational scales

Implicational scales are a form of scalogram analysis developed by Guttman (1941) for the measurement of social attitudes. In this analysis, the rows and columns of a qualita-

Tab. 39.4: Implicational scale

		Questions			
		A	B	C	D
Response patterns	1	1	1	1	1
	2	0	1	1	1
	3	0	0	1	1
	4	0	0	0	1
	5	0	0	0	0

tive data matrix can be arranged to fit a 'perfect scale' or a scalogram pattern. Table 39.4. shows five responses that are arranged from all 'yes' (all 1s) to all 'no' (all 0s) on four Questions. In this kind of arrangement, a 'yes' reply on Question *x* necessarily implies a 'yes' reply on all Questions to the right of Question *x* in the table. Similarly, a Response pattern *m* implies that all rows above Response pattern *m* have a greater number of 'yes' replies, while those below it have a smaller number of 'yes' replies.

Implicational scale analysis has a number of advantages. Arranging the data follows a fairly simple manipulation of rows and columns. The degree of fit of the model is measured by a statistical coefficient of scalability. Finally, the analysis provides a visual summary of the basic structure of the data under study. One of the pitfalls is how to deal with sparse data, that is, how to handle the presence of cells in the matrix that are empty due to a lack of information.

The first linguistic application of scalogram analysis was DeCamp's (1971) analysis of the Jamaican Creole continuum. Since then, this analytic tool has appeared often in language contact studies. An interesting example is Fasold's (1984: 231–239) analysis of a shift from Tiwa to English in the Tiwa Indian community in Taos, New Mexico. Tables 39.5a. and 39.5b. present a subset of the language contact data that Fasold reports. He finds that speakers over 40 years of age fit one implicational pattern perfectly, as shown in Table 39.5a, and that speakers between 16 and 35 years of age fit a different pattern, as shown in Table 39.5b. Note that a change in the position of the column labeled '2', which represents parents as interlocutors, provides an important clue to the nature of the language shift. This example shows that a single linear model would inadequately represent the language shift process for this community.

Manley (2014) uses scalogram analysis to examine cross-linguistic influence features in the speech of the members of two bilingual Quechua-Spanish communities in the city of Cuzco, Peru. Her study includes 31 phonetic, morphosyntactic, and calque features that reflect Quechua-into-Spanish influence. The data are arranged as a 70-participants by 31-features matrix, with presence/absence of a feature as the entry in each cell. Visually, the data matrix shows only a weak trend toward the scalogram pattern; that is, the data clearly do not fit a perfectly idealized implicational pattern. This observation is reflected by a statistical coefficient of scalability of 0.102 for these data, which falls below the threshold of 0.60 used by statisticians. This statistic indicates that it is not possible to claim scalability and unidimensionality of these data. From the perspective of quantitative analysis, the presence of a trend invites the researcher to mine the data further in order to isolate possible patterns. For example, a closer look at the participants-

Tab. 39.5: Tabulation of language-use patterns by members of a Tiwa Indian community (Fasold 1984: 235, 237)

a: Implicational pattern for language choice with seven interlocutors by speakers over 40 years of age

Speaker number	Speaker's age	1	2	3	4	5	6	7
1	65	T	T	T	T	T	TE	T
7	75	T	T	T	T	TE	TE	E
8	65	T	T	T	T	TE	E	E
18	41	T	T	T	TE	TE	TE	E
19	65	T	T	T	TE	TE	E	E
21	43	T	T	T	T	TE	TE	E
23	44	T	T	TE	TE	TE	E	E

b: Implicational pattern for language choice with seven interlocutors by speakers 16 to 35 years of age

Speaker number	Speaker's age	1	3	4	5	2	6	7
5	25	T	T	T	T	T	E	E
39	31	TE*	T	TE*	T	TE	TE	E
16	20	T	T	T	TE	TE	E	E
20	16	T	T	TE*	T	TE	E	E
24	17	T	TE	TE	T*	T*	E	E
32	17	TE	TE	TE	TE	TE	TE	E
33	18	TE	T*	TE	TE	TE	TE	E

Interlocutors are (1) grandparents, (2) parents, (3) Indian friends in the pueblo, (4) Indian friends in town, (5) brothers and sisters, (6) Indian friends when Anglo and Chicano friends are around, (7) Anglo and Chicano friends. Language-use patterns are T = Tiwa, E = English, TE = both Tiwa and English. An asterisk indicates cells that fail to conform to perfect scalability.

by-features matrix with frequencies in each cell − instead of the binary (presence/absence) entries − might provide a finer discrimination among participants and linguistic features.

5.2. Cluster analysis

Cluster analysis is used to subdivide a set of cases (individuals, objects) into groups − or clusters − with respect to qualitative or quantitative variables. The analysis produces hierarchical classifications by fusing cases into increasingly larger groups. A cluster of cases is a group whose behavior − as measured by the variables − is more like that of a centroid member of one particular cluster than it is of the centroid member of any other cluster. One of the main outputs is a tree diagram, known as a dendrogram, that visually

displays cases and clusters, and places them along a scale that indicates the distance or dissimilarity between cases and clusters.

Numerous algorithms can perform cluster analyses. Choosing a method includes considerations such as the distance measure and the clustering criteria. An important limitation of this type of analysis is the stability of the clusters. As McEntegart and Le Page (1982) point out, the stability of the clusters identified by an analysis should be established by the replication of the cluster solution with different clustering methods, by comparison with the results from other multivariate techniques (such as principal components analysis), and by qualitative evaluation.

Deumert (2004) applies hierarchical cluster analysis (Ward's method, average linkage, Euclidean distance measure) to data from the *Corpus of Cape Dutch Correspondence*. Her analysis groups 35 individual writers in this corpus on the basis of ten morpholexical and morphosyntactic variables that are diagnostic of Afrikaans as an independent linguistic system. Variables include: loss of neuter gender, loss of person and number distinctions, and use of Afrikaans pronoun variants. The analysis identifies three clusters that are interpreted as sociolinguistic varieties: (1) acrolectal Dutch, (2) mesolectal Cape Dutch vernacular, (3) the emerging standard of Afrikaans. Multidimensional scaling of the data confirms the three-cluster analysis and adds the nuance that the varieties form overlapping systems with regard to morphosyntactic variation.

Péronnet et al. (1998) apply cluster analysis in combination with correspondence analysis to study lexical variation in Acadian French spoken in eastern Canada. The data are from a dialect atlas of maritime vocabulary that shows contact between English, French, and Mi'kmaq (an Algonquian language). Individuals in the analysis are 54 fishermen from 18 different localities. The main clusters correspond closely to the geographic distribution of the fishermen's localities and show different degrees and types of lexical contact.

6. Hypothesis testing

A growing number of language contact studies make use of hypothesis testing. They analyze associations among variables in the context of a dependent-independent variable relationship, wherein one or more independent variables predict a dependent variable. Studies use both classical techniques, such as ANOVA and variable rule analysis, as well as recently developed multi-level modeling.

6.1. ANOVA

Analysis of variance (ANOVA) is used to test hypotheses about the effect of one or more categorical variables (often referred to as factor or independent variables), on a response or dependent variable (which is continuous). This statistical method tests differences between two or more means of the response variable, taking into consideration the amount of variance associated with each mean. ANOVA can capture different levels of a factor variable, and it can isolate interactions among two or more factor variables. This type of analysis is closely associated with experimental approaches in which the

factors under study are controlled. See Rietveld and van Hout (1993) for a detailed discussion of applications of ANOVA in linguistic studies; the authors pay special attention to the assumptions that underlie this model.

Applications of ANOVA to the heritage language data in Figure 39.2 test whether there is an effect of language and generation on VOT. Results (Hrycyna et al. 2011) show that there are significant effects of generation for Russian and for Ukrainian, but not for Italian. Follow-up post-hoc tests reveal that VOT is significantly higher in the Russian data for Generation 3 compared with Generation 1. In Ukrainian, Generation 5 has significantly higher VOTs than the other generations. The authors discuss these differences in terms of the structure of the three speech communities.

A study by Poplack, Sankoff, and Miller (1988) provides instructive remarks about some of the limitations of ANOVA. The authors investigate five factors that could influence the borrowing rate of English words into French: age, sex, social class, neighborhood, and proficiency in English. Application of ANOVA shows that social class membership and neighborhood are the main determiners of borrowing rate. However, when the authors try to study interactions between pairs of factors, they find empty cells (that is, missing data). Because these missing data create a situation where one of the assumptions of ANOVA is not met, the authors are forced to recode three of the factor variables in terms of simple binary distinctions in order to complete their analysis. This results in a loss of analyzable information.

6.2. Variable rule analysis

Variable rule analysis (see Sankoff [1988] for a general overview) builds a model that predicts a categorical response variable from several predictor (or independent) variables which are also categorical. The response variable is a linguistic feature that has two (or more) possible outcomes. The predictors are social and linguistic variables that are also categorical. The analysis evaluates the likelihood of each predictor variable co-occurring with the response variable. Using the strength of these relationships, the analysis establishes a hierarchy of the predictor variables and assigns probabilistic values to each level of a predictor variable to indicate its role in predicting the linguistic feature under study.

This technique is important because it is one of the few that tests hypotheses involving categorical data. In the terminology of general statistics, variable rule analysis is a version of logistic regression analysis. It is widely used in variationist sociolinguistics, where researchers make use of freely available software, such as the Goldvarb computer program (Sankoff, Tagliamonte, and Smith 2005). With recent advances in statistical theory (see below), there has been some updating of this approach.

Torres Cacoullos and Aaron (2003) study the patterning of determiners of lone English-origin nouns in New Mexican Spanish. They run variable rule analyses to examine the relation of syntactic, semantic, and discourse factors to these determiners in four separate datasets: single-occurrence English-origin nouns in Spanish discourse, established English-origin loanwords in Spanish discourse, Spanish nouns in unmixed Spanish discourse, and English nouns in unmixed English discourse. Comparison of the hierarchies of the predictor variables established by the analysis shows that lone English-origin nouns in Spanish behave like Spanish nouns in Spanish discourse. This result is dis-

cussed in terms of Poplack's nonce-borrowing hypothesis (see Poplack, Sankoff, and Miller 1988).

Delaere and De Sutter (2017) combine logistic regression analysis with correspondence analysis to study the variability of English loanword use in Belgian Dutch translations. Two main factors are analyzed: register and source language (non-translated Dutch, Dutch translated from French, and Dutch translated from English). Both statistical analyses show that register has a major effect: legal, political, and specialized texts have fewer loanwords than do journalistic texts. However, the analyses differ with respect to the role of source language. While correspondence analysis points to a register-by-source language interaction, this interaction could not be entered into the logistic regression model because of collinearity issues that yield an unstable model. The study underscores the complementarity between these two statistical techniques: one tests hypotheses, and the other takes a descriptive and exploratory approach. It also points to the advantages of using related techniques in tandem.

6.3. Mixed-effects models, random forests, and conditional inference trees

Recent advances in the field of statistics have proposed statistical tools that address difficulties which researchers encountered with classical approaches such as ANOVA and variable rule analysis. Some of these limitations are: interactions among the predictor variables; designs in which some cells have a very small number of tokens, or no tokens at all; the need to recode quantitative data into a small number of categories, which can lead to a loss of information (for example, converting a continuous variable, such as age, into two categories, e.g. below age 45 versus above age 45).

Mixed-effects modeling is one of these newer statistical tools. As in logistic regression analysis, mixed-effects models assess the importance of several factors in predicting a response variable that is categorical. An important feature of these models is that they distinguish between fixed and random predictor factors: fixed effects refer to factors that do not change from one study to another (such as sex, region), and random effects are factors that vary between studies (such as speakers, lexical items). By making this distinction, the models can assess the importance of random factors, such as individual speakers, in the prediction, a feature that is not available in logistic regression. In addition, both categorical and continuous predictor factors can be included in one analysis. Another technique, called random forests, gives information about the relative importance of the predictor variables. This is useful in research studies in which there is high collinearity among several predictor factors. A third technique, conditional inference trees, provides a flowchart decision tree that visualizes how different predictor factors operate together within a model. While these new analytical tools are powerful, rigid standards in their use are not always clear. See Tagliamonte and Baayen (2012) on the application of all three techniques to the analysis of sociolinguistic variation.

Adamou et al. (2016) explore the relationship between the percentage of borrowings in four minority Slavic languages (the dependent variable) and extralinguistic factors (the independent variables). The predictor factors include fixed factors (age, sex, language, text, and type [nouns or non-nouns]), and random factors (speaker and locality).

Mixed-effects modeling shows that, among the four languages, speakers of Molise Slavic produce the highest borrowing rates (see Figure 39.3. discussed earlier); it also reveals a number of complex interactions, such as age-by-type and sex-by-type, within each language. Random forests analysis shows that the most important factors in predicting the dependent variable are language and locality.

Szmrecsanyi et al. (2016) use random forests and conditional inference trees in tandem to analyze syntactic variation in four varieties of English (British, Canadian, Indian, and Singaporean). In a study of variation in particle placement (*he switched on the tube* versus *he switched the tube on*), they analyze the role of a number of syntactic and semantic factors. The structure of the conditional inference tree gives a visual representation of the complex interrelations among these effects, and random forest analysis suggests that word length and thematicity are the most important internal factors. From a language contact perspective, one of the interesting results is that significant cross-varietal differences occur mainly in short (three words or fewer) direct objects: Inner Circle (or L1) varieties (British, Canadian) have more split variants than do Outer Circle (or L2) varieties (Indian, Singaporean).

7. Software programs

SPSS (the Statistical Package for the Social Sciences) has been a standard research tool for many decades. This commercial software package uses a classic dropdown menu format and supports a wide range of statistical techniques. Eddington (2015) is an overview of applications of SPSS for many kinds of language data, and includes an informative statistical commentary on applications of advanced-level multivariate techniques.

Many researchers have recently adopted a software program called R (R Development Core Team). R is a programming environment that uses a command-line interface, allowing researchers to write their own code or modify existing code. R is a powerful statistical tool that is supported and regularly updated by many statisticians, produces impressive graphics, and is free of charge. Baayen (2008), Gries (2013), and Johnson (2008) introduce readers to R; they cover applications in many areas of linguistic inquiry. Software written in R for correspondence analysis is found in Greenacre (2017). Johnson (2009) presents Rbrul, a mixed-effects model version of variable rule analysis written in R. Larson-Hall (2016) is a guide to doing statistics using both R and SPSS; examples are drawn from second language research.

MATLAB (matrix laboratory) is a powerful, commercially available programming language. Among its applications are graphics, speech signal analysis, and statistical analysis. Finally, the Language Variation Suite (Scrivner and Díaz-Campos 2016) provides access to advanced statistical methods for researchers in linguistics who may have little knowledge of programming; this application is user-friendly and interactive.

Acknowledgments

Figures 39.2 and 39.3 were reprinted with permission from the authors.

8. References

Adamou, Evangelia, Walter Breu, Lenka Scholze & Rachel Xingjia Shen
2016 Borrowing and contact intensity: A corpus-driven approach from four Slavic minority languages. *Journal of Language Contact* 9(3). 513–542.

Baayen, R. Harald
2008 *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Bullock, Barbara E., Jacqueline Larsen Serigos & Almeida Jacqueline Toribio
2016 Stratification of English-language lone-word and multi-word material in Puerto Rican Spanish-language press outlets. In Rosa E. Guzzardo Tamargo, Catherine M. Mazak & M. Carmen Parafita Couto (eds.), *Spanish-English codeswitching in the Caribbean and the US*, 171–189. Amsterdam: John Benjamins.

Choi, Jinny K.
2005 Bilingualism in Paraguay: Forty years after Rubin's study. *Journal of Multilingual and Multicultural Development* 26(3). 233–248.

Cichocki, Wladyslaw
1986 *Linguistic applications of dual scaling in variation studies*. Toronto: University of Toronto dissertation.

Cichocki, Wladyslaw
1996 Data analysis. In Hans Goebl, Peter Nelde, Zdenek Stary & Wolfgang Wölck (eds.), *Contact linguistics: An international handbook of contemporary research*, 776–787. Berlin: De Gruyter.

Cichocki, Wladyslaw
2008 Retroflex realizations of Acadian French /R/: A dialectological perspective. *Toronto Working Papers in Linguistics* 27 (Special issue: *A Festschrift in honour of Jack Chambers)*. 17–30.

DeCamp, David
1971 Toward a generative analysis of a post-creole speech continuum. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 349–370. Cambridge: Cambridge University Press.

Delaere, Isabelle & Gert De Sutter
2017 Variability of English loanword use in Belgian Dutch translations: Measuring the effect of source language and register. In Gert De Sutter, Marie-Aude Lefer & Isabelle Delaere (eds.), *Empirical translation studies: New methodological and theoretical traditions*, 81–112. Berlin: De Gruyter.

Deumert, Ana
2004 *Language standardization and language change: The dynamics of Cape Dutch.* Amsterdam: John Benjamins.

Dorian, Nancy
1981 *Language death: The life cycle of a Scottish Gaelic dialect*. Philadelphia: University of Pennsylvania Press.

Eddington, David
2015 *Statistics for linguists: A step-by-step guide for novices*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Fasold, Ralph
1984 *The sociolinguistics of society.* Oxford: Oxford University Press.

Greenacre, Michael
2017 *Correspondence analysis in practice*. 3rd edn. Boca Raton, Florida: Chapman & Hall/CRC.

Gries, Stefan Th.
2013 *Statistics for linguistics with R*. Berlin & New York: De Gruyter Mouton.

Guttman, Louis
1941 The quantification of a class of attributes: A theory and method of scale construction. In Paul Horst, Paul Wallin, Louis Guttman & Ernest Watson Burgess (eds.), *The prediction of personal adjustment*, 321–347. New York: Social Science Research Council (US).
Hrycyna, Melania, Natalia Lapinskaya, Alexei Kotchetov & Naomi Nagy
2011 VOT drift in 3 generations of heritage language speakers in Toronto. *Canadian Acoustics* 39(3). 166–167.
Inoue, Fumio
1993 Word-accent change in progress in Japanese: Multivariate analysis of sociolinguistic groups. In André Crochetière, Jean-Claude Boulanger & Conrad Ouellon (eds.), *Proceedings of the XVth International Congress of Linguists* 3. 355–358. Sainte-Foy: Presses de l'Université Laval.
Johnson, Daniel E.
2009 Getting off the Goldvarb standard: Introducing Rbrul for mixed-effects variables rule analysis. *Language and Linguistics Compass* 3. 359–383.
Johnson, Keith
2008 *Quantitative methods in linguistics*. Malden: Blackwell.
Larson-Hall, Jenifer
2016 *A guide to doing statistics in second language research using SPSS and R*. London: Routledge.
Larson-Hall, Jenifer
2017 Moving beyond the bar plot and the line graph to create informative and attractive graphics. *The Modern Language Journal* 101(1). 244–270.
Le Page, Robert B., Pauline Christie, Baudoin Jurdant, A. J. Weekes & Andrée Tabouret-Keller
1974 Further report on the survey of multilingual communities. *Language in Society* 3. 1–32.
Ma, Roxana & Eleanor Herasimchuk
1971 Speech styles in Puerto-Rican bilinguals. In Joshua A. Fishman, Robert L. Cooper & Roxana Ma (eds.), *Bilingualism in the barrio*, 268–295. Bloomington: Indiana University.
Manley, Marilyn S.
2014 Measuring Quechua to Spanish cross-linguistic influence. In Jeffrey Connor-Linton & Luke Wander-Amoroso (eds.), *Measured language: Quantitative studies of acquisition, assessment, and variation*, 95–110. Washington DC: Georgetown University Press.
McEntegart, Damian & Robert H. Le Page
1982 An appraisal of the statistical techniques used in the "sociolinguistic study of multilingual communities". In Suzanne Romaine (ed.), *Sociolinguistic variation in speech communities*, 105–124. London: Edward Arnold.
Péronnet, Louise, Rose Mary Babitch, Wladyslaw Cichocki & Patrice Brasseur
1998 *Atlas linguistique du vocabulaire maritime acadien*. Québec: Presses de l'Université Laval.
Poplack, Shana, David Sankoff & Christopher Miller
1988 The social correlates of lexical borrowing and assimilation. *Linguistics* 26. 47–104.
Rasinger, Sebastian M.
2008 *Quantitative research in linguistics: An introduction*. London & New York: Continuum.
Rietveld, Toni & Roeland van Hout
1993 *Statistical techniques for the study of language and language behaviour*. Berlin: Mouton de Gruyter.
Rubin, Joan
1968 *National bilingualism in Paraguay.* The Hague: Mouton.
Sankoff, David
1988 Variable rules. In Ulrich Ammon, Norbert Dittmar & Klaus J. Mattheier (eds.), *Sociolinguistics: An international handbook of the science of language and society*, vol. 2, 984–997. Berlin: Walter de Gruyter.

Sankoff, David, Sali A. Tagliamonte & Eric Smith
2005 Goldvarb X: A multivariate analysis application [computer program]. Toronto & Ottawa: Department of Linguistics, University of Toronto & Department of Mathematics, University of Ottawa.
Scrivner, Olga & Manuel Díaz-Campos
2016 Language variation suite: A theoretical and methodological contribution for linguistic data analysis. *Proceedings of the Linguistic Society of America* 1. 1–15.
Szmrecsanyi, Benedikt, Jason Grafmiller, Benedikt Heller & Melanie Röthlisberger
2016 Around the world in three alternations: Modeling syntactic variation in varieties of English. *English World-Wide* 37(2). 109–137.
Tagliamonte, Sali A. & R. Harald Baayen
2012 Models, forests and trees of York English: *was*/*were* variation as a case study for statistical practice. *Language Variation and Change* 24(2). 135–178.
Torres Cacoullos, Rena & Jessie Elana Aaron
2003 Determiner variation with English-origin nouns in New Mexican Spanish: Borrowing bare forms. *University of Pennsylvania Working Papers in Linguistics* 9(2). 159–172.
Woods, Anthony, Paul Fletcher & Arthur Hughes
1986 *Statistics in language studies*. Cambridge: Cambridge University Press.

Wladyslaw Cichocki, Fredericton (Canada)

40. Qualitative data elicitation and analysis

1. Introduction: what are qualitative methods?
2. The development of sociological and interactionist approaches to multilingual language use
3. Qualitative fieldwork methods
4. Qualitative methods for data analysis
5. Conclusion
6. References

1. Introduction: what are qualitative methods?

Qualitative approaches have been commonplace in investigations of language contact since the very beginning of the field, as well as in analyses of structural variation and change. The work of scholars like Fishman, Hymes, and Gumperz from the 1960s onwards has defined qualitative methods in the investigation of multilingual language use as we think of them today, i.e. as sociologically and interactionist oriented approaches. We offer an overview of the development of these approaches to language contact (Section 2), present contemporary qualitative field methods and data analyses, and exemplify their use (Sections 3 and 4).

Qualitative methods usually involve an interpretative argument, and focus on the individual and the particular (Benson 2013; Taylor, Bogdan, and DeVault 2015). As a consequence of this interest in the specific rather that the general, instead of a focus on larger populations, we frequently find case studies, which means that findings are often seen as limited in terms of generalizabilty. Despite this, qualitative methodology can also be used

https://doi.org/10.1515/9783110435351-040

in studies aiming at investigating the 'typical', and in these cases, generalizability is then argued for.

While the distinction between qualitative and quantitative methods may be a helpful tool for categorizing research methods, much recent research within the fields of contact linguistics, sociolinguistics, and dialectology combines approaches, or uses approaches traditionally viewed as belonging to one paradigm to shed light on questions typically asked within the other.

2. The development of sociological and interactionist approaches to multilingual language use

We first highlight a number of studies that, in hindsight, represent milestones. Early studies of bilingual situations in the 1960s and 1970s mainly addressed the functional division of two or more varieties in multilingual communities. Here, the work of Fishman (1926–2015) and Gumperz (1922–2013) was methodologically innovative and had a huge impact on sociologically oriented and interactionist studies of multilingual language use. Fishman (1967, 1972) extended the notion of 'diglossia', introduced by Ferguson (1959), to cover the functional differentiation of stylistic registers, dialects, and separate languages, and through his work, the concept of linguistic domains became widely known (Fishman 1964, 1972). He operationalized it with the famous dictum, "Who speaks what language to whom and when?" (Fishman 1965). Methodologically, Fishman's studies were based on participant observation of social networks and individual participants. He employed observation as a method in its own right, but also as a means of hypothesis building and as a starting point for further qualitative and quantitative data collection, i.e. as a foundation for large-scale social investigations based on censuses and questionnaires (Fishman 1971). He advocated linking the observation of language use with self-report methods, such as via subsequent interviews of speakers or the use of attitude scales and questionnaires, thus improving the reliability and validity of the results (Fishman 1969; see also García and Schiffman 2006: 14–18). Interweaving micro- and macro-level parameters by combining different methodological approaches played an essential role in Fishman's work, as well as in much variationist contact linguistics. Another methodological contribution promoted by Fishman, one which has become the norm today, is the axiom that researchers should seek to establish an insider's view on the social group under study in order to gain a deeper understanding of how language use relates to behavioral and/or attitudinal patterns.

Gumperz contributed to sociologically oriented studies on multilingual language use, often referred to under the heading of 'interactional sociolinguistics' (see Section 4 below). Methodologically, he called for in-depth analyses of selected instances of recorded verbal interaction in order to pinpoint how speakers create, define, and reproduce the symbolic value of varieties, as well as to demonstrate how specific communicative acts become meaningful in specific settings. He advocated for the long-term engagement of the researcher with a particular speech community in order to gather information on language use and attitudes through participant observation, and also called for further ethnographic investigation into other social factors, such as kinship, ethnicity, land tenure, and so on. Like Fishman, he employed repeated participant questioning about speak-

ers' linguistic choices. Based on analyses of recorded bilingual speech, Gumperz developed the concept of 'contextualization cues', describing how switching between two (or more) varieties may add an extra layer of meaning. Furthermore, he introduced the distinction between situational and metaphorical codeswitching: situational code-switching should be predictable on the basis of changes in external factors (such as the arrival of new participants in the interaction), whereas metaphorical code-switching denotes an unexpected switch, which is not related to external factors, but serves internal interactional functions (Blom and Gumperz 1972; Gumperz 1982: 62).

In the 1980s the act of switching itself became an important focus for studies of language contact. Switching between languages or varieties was established as a social cue, a communicative signal to one's co-participants. Auer (1984), extending Gumperz' notion of contextualization cues, applied a Conversation Analysis (CA) approach to German-Italian bilingual data, hereby adding the CA approach to the field of interactionist analyses of multilingual practices, choices, and patterns. A CA approach examines speakers' language choices within the context of a particular conversation; it does not take external social factors into direct account, unless such factors are made relevant in the conversation by the participants themselves. Other key studies that combined a CA approach with ethnographic observation and the analysis of (recorded) speech are Li Wei's (1994) study of three generations of Chinese-British families, and Li Wei, Milroy, and Pong's (1992) study of a Chinese community in northeastern England.

The last two decades have seen a rise in anthropologically inspired studies of small multilingual groups and individuals which focus on situational linguistic positioning in a complex, late modern world, often in an urban context, and, again, taking into account the wider societal and institutional structures (e.g. class stratification and ethnicity) that influence the communicative setting (e.g. Keim 2007; Madsen 2015; Rampton 2006). With respect to methodology, studies still largely rely on the qualitative methods developed in the early years of sociolinguistic research on bilingual language use, with innovations primarily centered on the development and use of multi-methodological approaches. Rampton (1995) investigated language use in migrant youth groups in a multilingual neighborhood in southern England, providing an excellent example of this multi-methodological approach. Another example of a combined methodological approach is a study of Moroccan boys in Antwerp by Jaspers (2006). These studies rely on ethnographic participant observations over long stretches of time in different settings, recordings of both naturally occurring interactions and interviews, and fine-grained analyses of speakers' language choice and use in the recordings.

3. Qualitative fieldwork methods

We focus here on the following methods: *ethnographic* methods to collect *data on communicative practice*, including observation and interview; and methods to collect data for analyzing *discourses and norms*. The different methods are not clearly delimited, but traditionally, different types of research questions have tended to entail different methods of data collection.

The aim of ethnography is to understand social meaning-making. This can be theoretically conceived of in different ways. Malinowski wrote already in 1922 that the final

aim of ethnography is to: "grasp the native's point of view, his relation to life, to realize his vision of his world" (Malinowski 1922: 25). As mentioned above, this perspective was crucial to Fishman's work on language contact. However, social constructionist approaches have questioned and criticized this aim and its underlying assumptions (Clifford and Marcus 1986). One line of criticism claims that talk about a 'participant perspective' presupposes that there *is* such a relatively stable perspective, whereas meaning-making in interaction is complex and multifaceted, with so-called perspectives varying from situation to situation, and even from moment to moment (Blommaert 2013). As a result, exact descriptions of participant perspectives in ethnographic studies must be regarded as an illusion. What we *can* describe is social behavior, including the use of language. This general debate within the social sciences has led to a more critical perspective on ethnographic methods, including in language contact research (Heller 2011).

The first step in preparing fieldwork that focuses on understanding how and why linguistic variation is used will usually be to delimit the group under investigation. Sometimes this selection can consist of just one person (Coupland 1980), but most often, the selection will be based on prior knowledge of the community and involve a group of people with preexisting mutual interactions, such as a school class, the work place, a family, and so on (Li Wei 1994; Otsuji and Pennycook 2010; Rampton 1995). Studies may also focus on people selected out of a broader field, with no prior social relationships with one another.

3.1. Observations

Observations of linguistic practice across many contexts form the basis of ethnographic research within the field (cf. the work by Fishman, Gumperz, and Rampton in Section 2). Usually ethnographic observation takes place over a longer period of time, with the ethnographer observing, for instance, on a daily basis in a specific field.

It is crucial to carry out these observations in a systematic and well-documented fashion in order to produce reliable data. Documentation usually involves field notes, diaries, drawings, and perhaps also photos and videos. The aim is to document what we as researchers learn from observations in the field. Here-and-now documentation allows us to discover patterns of linguistic behavior. The collected data in many cases also document how the researcher gains an increasing understanding of the field over time, and thereby becomes increasingly able to connect small linguistic details to larger frames of social understanding (for introductions to methods of ethnographic observations in multilingual settings, see Heller, Pietikäinen, and Pujolar [2017] and Copland and Creese [2015]). Detailed documentation of the data collection process is decisive in order to enable others to discuss the findings in an informed way.

3.2. Recordings

Another central type of data is recordings of interaction. For language contact, we are typically interested in knowing how language is perceived and made socially meaningful on a moment-to-moment basis by people's linguistic practices − how and why do they

vary their use of specific linguistic material? Data for such analyses consist of audio- or video-recorded conversations, which can be elicited with more or less control on the part of the researcher. In some types of studies, the researcher will have the informants wear lapel microphones for longer periods as they carry out their daily routines (Schøning and Møller 2009); in others, it will be more informative to set up specific events for recordings. Recordings of everyday activities are generally easier to obtain if a relationship between the researcher and participant(s) is established in advance − for example, through ethnographic fieldwork. Studies of contact phenomena often focus on specific settings (Pennycook and Otsuji 2015) and/or a specific group of people (Madsen et al. 2016). The strategy for recording should of course reflect the focus of interest. In recordings of everyday activities, it is particularly important to consider ethical issues, and we advise that informed consent be obtained from all participants (for a discussion of informed consent in linguistic inquiry, see Mortensen 2015).

3.3. Interviews

Interviews including one or more individuals provide the researcher with information on how participants talk about language (or other subjects), and thus offer data for discourse or interaction analysis (see Section 4, below; see also Chapter 49). This type of data is valuable as a source of information about how different languages are ideologically linked to different domains (see again Fishman and Gumperz, Section 2), or how other norms of use are produced in discourse. Interviews for ethnographic and/or sociolinguistic purposes are often semi-structured (Kvale 1996), meaning that there is a structure for the interview prepared by the researcher which the conversation should more or less follow, with other topics and questions included depending on how the conversation unfolds. Interviews might also be used to obtain participants' comments on previously collected data (Rampton 2006). As such, interviewing is an important method for eliciting data on ideology and norms, but interviews should not be seen as providing data on everyday linguistic practices. To understand how language is used in everyday practice, we need other types of data, such as self-recordings and ethnographic observations.

3.4. Written texts on social media

The growing importance of social media in contemporary society means that in recent years, written material has become an important source of data in new ways. Communication on social media offers a completely new type of written data, namely written interactional data that happens very quickly in real time. This type of data can be difficult to access and collect for ethical and practical reasons (see Boellstorff et al. [2012] and D'Arcy and Young [2012]); communication on social media nevertheless forms an important part of many people's everyday interactions, and should consequently be viewed as an integral part of social life (see Androutsopoulos [2007, 2013] on multilingual practices in social media).

3.5. Critical discourse analysis

Critical discourse analysis of written documents is central to research in language contact situations. Texts such as educational policies or institutional language policies constitute important exercises of power, whereby institutions seek to influence individuals' concrete linguistic practices in specific ways. Thorough analyses of such texts are crucial to increase our understanding of the workings and roles of different languages within a specific society. Fieldwork of this sort includes collecting material representing language policies from public domains, or from private companies and organizations. This entails gaining access to specific archives or documents, and even though an official language policy is most often made public in order to function, discussions, debates, and processes of negotiation leading up to the final version of the policy may not be easily accessible. However, these can be important data for the researcher, showing how different discourses are present in arguments about regulation of language use. Political documents, reports, and laws should be approached through use of qualitative methods, since a detailed discourse analysis will be able to account for the ideologies and meanings associated with different languages in such texts. Within the area of language contact, studies belonging to this field of inquiry include studies of national language policy in places like China (Zhou and Sun 2004), Catalonia (Piulats 2007), or Kenya (Kibui 2014; Ogechi 2003). Many include analyses of both institutional documents and situated linguistic practices, showing how powerful discourses are produced, reproduced, and contested through the dialectical relationship between top-down regulations and bottom-up practices (see also Chapter 51).

Thus far, we have outlined different methods for collecting data in the field. Which data to collect, and how to do so, is ultimately tied to the attendant research questions, and these must be clearly stated from the beginning of the data collection phase in order to ensure a strong connection between research data, and the possible conclusions based thereon. In Section 4, we turn to the next methodological stage: data analysis.

4. Qualitative methods for data analysis

In contact linguistics, qualitative data analysis serves to investigate how people perceive, manage, and exploit linguistic contact phenomena in their daily social lives. As we illustrate in this section, it enables researchers to explore how people organize situational use of different linguistic resources, how linguistic diversity enters into, and is exploited to create, situated understanding, and how people negotiate social relations in connection to perceived group boundaries by using language perceived as 'belonging' to other groups (Rampton 1995). In other words, qualitative methods applied to contact phenomena investigate how people engage with linguistic diversity in their everyday lives. We sketched out some methods for collecting data to gain such knowledge above; in this section, we focus on ways of analyzing these data.

The aim here is not to provide a guide for how to use qualitative analytical methods, but rather to introduce central analytical approaches and illustrate where such analytical work can take us. As opposed to variationist methods that are hypothesis-driven, qualitative methods permit open-ended research questions. Often, not least in ethnographic

studies, these questions will change along the way, as the researchers gain increasing knowledge about the field. Qualitative studies further differ from many variationist studies in that the former are not dependent on predefined categories. Therefore, qualitative approaches to linguistic contact phenomena have the potential to challenge fixed or static perceptions of language categories. Furthermore, these studies may provide new knowledge about people's perceptions of (new) ways of speaking, how such perceptions become valorized, and how groups of speakers identify with the use of specific linguistic contact phenomena by, for example, claiming to be authentic users of a particular linguistic practice.

Qualitative analysis, at the most basic level, addresses social phenomena in situated settings and emphasizes the perspectives and views of the participants (Pope and Mays 1995). In work with linguistic data, this means that in practice, the analytical process involves zooming in and out between the fine-grained details and the larger picture seen in the data. In the following, we briefly sketch two central frameworks for analyzing linguistic contact phenomena: interactional sociolinguistics and grounded theory.

4.1. Interactional sociolinguistics

In *interactional sociolinguistics* (Gumperz 1982; Rampton 2017), a key aim is to explore how variability in language interrelates with situated understanding. Traditionally, the method was applied to analyses of face-to-face interactions, but many of the principles are applicable to other sorts of interaction as well, such as, for example, exchanges on social media. The analytical strategy is to analyze how interlocutors employ and react to communicative resources. Interaction is viewed as an activity wherein the participants are constantly engaged in making sense of each other's communicative signs (this is called *inferencing*), and providing cues on a moment-to-moment basis that guide one toward how these signs should be understood (Gumperz' *contextualization cues*; cf. Section 2).

A central assumption is that interlocutors engaged in the activity of inferencing draw on their sociolinguistic knowledge of the interpretation of cues, and their associations of certain ways of speaking with groups of speakers, types of behaviors, and so on. The phenomenon of *crossing* (Rampton 1995) is a good example of this. The term points to moments in interactions where speakers use languages which are generally thought to belong to other (ethnic) groups of speakers. An example is when young people with an Anglo or Afro-Caribbean background living in London use bits of Panjabi in the presence of peers with an Indian or Pakistani background (Rampton 1995: 4), which may connote a "sense of movement across social or ethnic boundaries" (Rampton 1995: 280). However, in order for this activity to be made meaningful in a specific interaction, the participants must have some shared knowledge which allows them to recognize particular clusters of features as associated with certain ways of speaking. Furthermore, they must have shared knowledge about who the stereotypical users of these ways of speaking could be, and shared normative views about 'legitimate' and 'illegitimate' users. This illustrates how perceptions of linguistic and cultural differences may enter into situated understanding (Gumperz 1992). With its focus on how perceptions of ways of speaking become meaningful in interaction, interactional sociolinguistics links up with contact linguistics.

The specific analytical work of interactional sociolinguistics involves thorough listening of sound (or video) recordings and careful transcriptions − either of the whole material, or of selected key moments (for a more detailed introduction, see Jaspers [2011] and Rampton [2006]). Selecting these moments is part of the analytical process, and the reasons for choosing them must always be explained in detail. Interactional sociolinguistics is related to Conversation Analysis (see above), in the sense that it also examines turn-taking patterns, and a central analytical step is to consider how the participants treat each other's utterances. However, a central aim for interactional sociolinguistics is to understand how situated language use interplays with broader frames of social understanding (see also Section 3.1). If the researcher aims to understand meaning-making in interaction, knowledge about the social history of the participants, of how they organize language, and more generally, of the indexical potential (Agha 2007) of the language used in the exchange may become relevant as well. Such knowledge can also be gained from interviews and ethnographic observations, as mentioned in Section 2 and 3 above, and illustrated in the examples below.

4.2. Grounded theory

Working with rich datasets, such as interviews and ethnographic field notes, it is tempting to 'cherry-pick' one or two quotes that happen to support a specific interpretation, and discard the rest of the content. This, of course, may lead to biased results, which end up saying more about the researcher's preconceived notions of language than about the participants' perspectives on contact phenomena. One way to avoid this is to work from the perspective of *grounded theory* (Glaser and Strauss 1967; Strauss and Corbin 1990). Charmaz (2006: 2) describes the method of grounded theory as a tool for empirical research which provides "systematic, yet flexible guidelines for collecting and analyzing qualitative data to construct theories 'grounded' in the data themselves". A central activity in this approach is several rounds of coding, starting with an open coding procedure which essentially answers the question 'what is going on here?' in all parts of the dataset. This provides an overview, from the perspective of the researcher, of what the dataset contains. The subsequent rounds of coding focus on how the concepts from the initial coding cluster together and are interrelated. This procedure allows the researcher to categorize groups of codes, and in turn, to formulate inductively derived theories (for a detailed introduction, see Charmaz [2006] and Strauss and Corbin [1990]) which can be related to, and possibly challenge, or be challenged by, preexisting theories about human phenomena.

The method of grounded theory invites researchers to consider and make sense of all parts of a dataset, and to consider and document all steps in the analytical process toward an interpretation. It also illustrates three additional general principles for qualitative work with rich datasets: a) *openness* to anything the data may contain, even if it is unexpected by the researcher; b) *systematicity* in the process of surveying the data; and c) *grounding*, in the sense that the researcher should always be able to point to specific features in the data that led to their analytical conclusions, and at the same time, be sure that other portions of the data do not undermine their conclusions.

4.3. Sample analysis

The approaches of interactional sociolinguistics and grounded theory described here can be fruitfully combined, if we seek to understand linguistic contact phenomena as social acts. To illustrate this, we show an example from an ethnographically based study from Copenhagen (analyzed in further detail in Stæhr [2015]).



Extract 1: Facebook, Bashaar on Isaam's wall (Stæhr 2015: 35)

The boy, Bashaar, congratulates another boy, Isaam, on the latter's Facebook wall. This is done with two words, an emoticon, and punctuation. The first word, *Habiibi*, is generally associated with Arabic, and can be translated as 'my love' or 'my friend'. The second word, *Tjellukka*, is generally associated with Danish, and can be translated as 'congratulations'; it is spelled by Bashaar in a way that differs significantly from the standard orthographic spelling *tillykke*. In this sense, we identify in this piece of data a 'contact phenomenon' involving Arabic and Danish.

The exchange is part of a large dataset comprising observations of on- and offline behavior, sound recordings, interviews, written material, and so on (Madsen et al. 2016). From several other analyses of interaction based on these data, we know that Isaam, Bashaar, and several other participants (all with a range of different linguistic backgrounds) frequently use features associated with a range of languages, such as Arabic, Turkish, Kurdish, and English, alongside features associated with Danish. In interviews, the participants display a large degree of attention to this phenomenon, and describe it as a component of what they label 'street language' (Madsen 2013).

The method of grounded theory provides a tool to process, and obtain an overview of, the many different data sources. To exemplify this, let us examine an extract from a group interview, where two girls who are part of the same project, Lamis and Selma, participate (Madsen, Spindler Møller, and Jørgensen2010: 93). The extract follows a sequence where the participants described ways of speaking they had come across in their daily life. The field worker attempts to elicit details about how the participants perceive these ways of speaking:

1 Int: but street language is it then
2 Selma: (.) it is like [mixed]
3 Lamis: [like that]
4 Selma: Turkish Arabic Danish
5 Lamis: yes like that
(.) = micro pause; [ ] = overlapping speech; names anonymized

Extract 2: Group interview

In essence, this piece of data links the label of 'street language' to a specific linguistic behavior of juxtaposing features associated with different languages. As it turned out, this was one out of many pieces that provided an understanding of the phenomenon 'street language'. The extract may be coded under some initial concepts, such as 'street language' and 'language mixing', and in the next round of coding, under 'metalinguistic description'. Processing data through coding like this provides an overview of, as well as detailed knowledge about, what we can learn from the data.

Returning to Extract 1, observations among the project participants confirm that the term *habibi* clusters together with other linguistic features associated with 'street language'. In this way, zooming out and taking a broader look at our data provides a new, more participant-oriented understanding of Extract 1, whereby we can now also view the two words *tjellukka* and *habiibi* as associated with the same way of speaking, namely with what is described by the participants as 'street language'. In this way, the example illustrates a central point in connection to contact linguistics: qualitative approaches have the potential to change our perception of what type of contact phenomenon we are dealing with. Without any attempt to view the interaction from the participants' perspective, researchers run the risk of seeing contact between different 'languages' (or other ways of speaking) in situations where the participants are experiencing something else. We do not claim that these linguistic features are no longer associated with Danish or with Arabic, or that the participants would be unable to identify them as such if we were to ask them. Rather, what we are observing here may be described as developments in *orders of indexicalities* (Silverstein 2003), wherein the indexical references to some of the participants' individual linguistic backgrounds are re-contextualized to signal 'belonging' within their peer group.

In our observations of the participants in this school more generally, a frequently observed and highly salient feature is the palatalization of /t/ (Hyttel-Sørensen 2017; Stæhr 2015). The participants use this feature alongside other features associated with street language, and Isaam (the recipient of the birthday greeting) specifically uses this feature to imitate street language in an interview with a fieldworker. This information put the spelling of *tjellukka* in Extract 1 in a different light, as the *tje*-spelling can be interpreted as an orthographic representation of the palatalized pronunciation of /t/ as [tʲ]. A review of changes in Bashaar's orthographic production in exchanges on Facebook over time further supported this interpretation: Bashaar used the same spelling feature (replacing *t* with *tj*) on other occasions as well. The knowledge that this feature is part of an in-group language shared by Isaam and Bashaar allows us to argue that this way of writing functions as a contextualization cue in the message − which further underscores the in-group nature of the greeting. Additionally, we know from the ethnography that Bashaar and Isaam were very close friends at the time of the study.

We have illustrated here the interplay between different methods in qualitative analysis. The example drew on micro-analyses which invited us to revisit larger parts of the data to look for patterns, after which we returned to the micro-analyses and understood them in new ways. The interviews about language also provided important information − not about how our participants interacted, but about how they perceived different ways of speaking and which features they stereotypically associated with each other.

More broadly, we have illustrated how qualitative analyses can help illuminate social dimensions of contact phenomena, with a focus on how people use and perceive these

phenomena. In this sense, qualitative studies may add nuance to researchers' perceptions of multilingual language use.

5. Conclusion

As should be evident from our brief account of the development of sociologically and interactionist oriented approaches in contact linguistics, these methods have been part of the study of language contact, in particular since the 1960s.

The field of qualitative methods is extremely broad. We have stressed the varied methodological approaches, and provided examples of the most commonly used methods for data collection within contact linguistics. For data analysis, we focused on two general approaches, interactional sociolinguistics and grounded theory, which were applied in an illustrative example. Our main objective has been to show the complexity of qualitative methodologies, and to illustrate that qualitative approaches to fieldwork and data analysis can contribute something different as compared to results gained from quantitative methods. Sometimes, the findings are of a completely different nature, as illustrated in Section 4. Qualitative methods are indispensable in any field of linguistics where language is seen as a social phenomenon.

6. References

Agha, Asif
2007 *Language and social relations*. Cambridge: Cambridge University Press.

Androutsopoulos, Jannis
2007 Language choice and code-switching in German-based diasporic web forums. In Brenda Danet B. & Susan C. Herring (eds.), *The multilingual internet: Language, culture, and communication online*, 340–361. Oxford: Oxford University Press.

Androutsopoulos, Jannis
2013 Networked multilingualism: Some language practices on Facebook and their implications. *International Journal of Bilingualism* 19(2). 185–205.

Auer, Peter
1984 *Bilingual conversation*. Amsterdam: Benjamins.

Benson, Philip
2013 Qualitative methods: Overview. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics*. London: Blackwell. DOI: 10.1002/9781405198431.wbeal0977

Blommaert, Jan
2013 *Ethnography, superdiversity and linguistic landscapes: Chronicles of complexity*. Bristol: Multilingual Matters.

Blommaert, Jan & Jenny-Louise van der Aa
2015 Ethnographic monitoring and the study of complexity. *Tilburg Papers in Culture Studies* 123. Tilburg: Tilburg University.

Blom, Jan-Petter & John J. Gumperz
1972 Social meaning in linguistic structures: Code switching in Northern Norway. In John J. Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics: The ethnography of communication*, 407–434. New York: Holt, Rinehart and Winston.

Boellstorff, Tom, Bonnie Nardi, Celia Pearce & T. L. Taylor
2012 *Ethnography and virtual worlds: A handbook of method*. Princeton: Princeton University Press.
Charmaz, Kathy
2006 *Constructing grounded theory: A practical guide through qualitative analysis.* London: Sage.
Clifford, James & George E. Marcus (eds.)
1986 *Writing culture: The poetics and politics of ethnography.* Berkeley: University of California Press.
Copland, Fiona & Angela Creese
2015 *Linguistic ethnography.* Los Angeles: Sage.
Coupland, Nikolas
1980 Style-shifting in a Cardiff work-setting. *Language in Society* 9(1). 1–12.
D'Arcy, Alexandra & Taylor Marie Young
2012 Ethics and social media: Implications for sociolinguistics in the networked public. *Journal of Sociolinguistics* 16(4). 532–546.
Evans, Stephen
2000 Hong Kong's new English language policy in education. *World Englishes* 19(2). 185–204.
Ferguson, Charles A.
1959 Diglossia. *Word* 15. 325–340.
Fishman, Joshua A.
1964 Language maintenance and language shift as a field of inquiry. *Linguistics* 2(9). 32–70.
Fishman, Joshua A.
1965 Who speaks what language to whom and when? The analysis of multilingual settings. *La Linguistique* 1. 67–88.
Fishman, Joshua A.
1967 Bilingualism with and without diglossia, diglossia with and without bilingualism. *The Journal of Social Issues* 23. 29–38.
Fishman, Joshua A.
1969 A sociolinguistic census of a bilingual neighborhood. *American Journal of Sociology* 75(3). 323–339.
Fishman, Joshua A.
1971 *Bilingualism in the Barrio*. Bloomington: Indiana University Press.
Fishman, Joshua A.
1972 Domains and the relationship between macro- and microsociolinguistics. In John J. Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics: The ethnography of communication*, 435–453. New York: Holt, Rinehart and Winston.
Gal, Susan
1979 *Language shift: Social determinants of linguistic change in bilingual Austria*. San Francisco: Academic Press.
García, Ofelia & Harold Schiffman
2006 Fishmanian sociolinguistics: 1949 to the present. In Ofelia García, Rakhmiel Peltz & Harold Schiffman (eds.), *Language loyalty, continuity and change: Joshua A. Fishman's contributions to international sociolinguistics*, 3–68. Clevedon: Multilingual matters.
Glaser, Barney & Anselm Strauss
1967 *The discovery of grounded theory.* London: Weidenfeld and Nicholson.
Gumperz, John J.
1982 *Discourse strategies*. Cambridge: Cambridge University Press.
Gumperz, John J.
1992 Contextualization revisited. In Peter Auer & Aldo Di Luzio (eds.), *The contextualization of language*, 39–54. Amsterdam: John Benjamins.

Heller, Monica
2011 *Paths to post-nationalism: A critical ethnography of language and identity.* Oxford: Oxford University Press.
Heller, Monica, Sari Pietikäinen & Joan Pujolar
2017 *Critical sociolinguistic research methods: Studying language issues that matter*. London: Routledge.
Hyttel-Sørensen, Liva
2017 *"Gangster" or "wannabe": Experimental and ethnographic approaches to a contemporary urban vernacular in Copenhagen*. Copenhagen: University of Copenhagen dissertation.
Jaspers, Jürgen
2006 Stylizing Standard Dutch by Morrocan boys in Antwerp. *Linguistics and Education* 17. 131–156.
Jaspers, Jürgen
2011 Interactional sociolinguistics and discourse analysis. In James Paul Gee & Michael Handford (eds.), *The Routledge handbook of discourse analysis*, 135–146. London: Routledge.
Keim, Inken
2007 Socio-cultural identity, communicative style, and their change over time: A case study of a group of German-Turkish girls in Mannheim/Germany. In Peter Auer (ed.), *Style and social identities: Alternative approaches to linguistic heterogeneity*, 155–186. Berlin: Mouton de Gruyter.
Kibui, Agnes W.
2014 Language policy in Kenya and the New Constitution for Vision 2030. *International Journal of Educational Science and Research* 4(5). 89–98.
Kvale, Steinar
1996 *InterViews: An introduction to qualitative research Interviews*. London: Sage.
Li Wei, Lesley Milroy & Sin Ching Pong
1992 A two-step sociolinguistic analysis of code-switching and language. *International Journal of Applied Linguistics* 2(1). 63–86.
Li Wei
1994 *Three Generations, two Languages, one family*: *Language choice and language shift in a Chinese community in Britain*. Clevedon: Multilingual Matters.
Madsen, Lian Malai
2013 'High' and 'low' in urban Danish speech styles. *Language in Society* 42(2). 115–138.
Madsen, Lian Malai
2015 *Fighters, girls, and other identities: Sociolinguistics in a martial arts club.* Bristol: Multilingual Matters.
Madsen, Lian Malai, Janus Spindler Møller & J. Normann Jørgensen
2010 "Street language" and "integrated": Language use and enregisterment among late modern urban girls. In Lian Malai Madsen, Janus Spindler Møller & J. Normann Jørgensen (eds.), *Ideological constructions and enregisterment of linguistic youth styles: Copenhagen studies in bilingualism*, vol. 55, 81–113. Copenhagen: University of Copenhagen, Faculty of Humanities.
Madsen, Lian Malai, Martha Sif Karrebæk & Janus Spindler Møller
2016 *Everyday languaging: Collaborative research on the language of children and youth*. Berlin: Mouton de Gruyter.
Malinowski, Bronislaw
1922 *Argonauts of the Western Pacific.* London: Routledge.
Mortensen, Kristine Køhler
2015 Informed consent in the field of language and sexuality: The case of online dating research. *Journal of Language and Sexuality* 4(1). 1–29.

Ogechi, Nathan Oyori
2003 On language rights in Kenya. *Nordic Journal of African Studies* 12(3). 277–295.
O'Reilly, Karen
2011 *Ethnographic methods*, 2nd edn. London: Routledge.
Otsuji, Emi & Alistair Pennycook
2010 Metrolingualism: Fixity, fluidity, and language in flux. *International Journal of Multilingualism* 7(3). 240–254.
Pennycook, Alastair & Emi Otsuji
2015 *Metrolingualism: Language in the city*. London: Routledge.
Piulats, Maria Corominas
2007 Media policy and language policy in Catalonia. In Michael J. Cormack & Niamh Hourigan (eds.), *Minority language media: Concepts, critiques and case studies*, 168–187. Clevedon: Multilingual Matters.
Pope, Catherine & Nick Mays
1995 Reaching the parts other methods cannot reach: An introduction to qualitative methods in health and health services research. *BMJ* 1(311). 42–45.
Rampton, Ben
1995 *Crossing: Language and ethnicity among adolescents.* London: Longman.
Rampton, Ben
2006 *Language in late modernity: Interaction in an urban school*. Cambridge: Cambridge University Press.
Rampton, Ben
2017 Interactional sociolinguistics. *Working Papers in Urban Language & Literacies* 205. https://www.academia.edu/30796363/WP205_Rampton_2017._Interactional_Sociolinguistics (accessed on 3 January 2019).
Schøning, Signe & Janus S. Møller
2009 Self-recordings as a social activity. *Nordic Journal of Linguistics* 32(2). 245–269.
Silverstein, Michael
2003 Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23. 193–229.
Stæhr, Andreas
2015 Reflexivity in Facebook interaction: Enregisterment across written and spoken language practices. *Discourse, Context & Media* 8. 30–45.
Strauss, Anselm L. & Juliet M. Corbin
1990 *Basics of qualitative research*. London: Sage.
Taylor, Steven J., Robert Bogdan & Marjorie DeVault
2015 *Introduction to qualitative research methods: A guidebook and resource.* London: Wiley.
Tollefson, James W. (ed.)
2002 *Language policies in education: Critical issues*. London: Lawrence Erlbaum Ass.
Zhou, Minglang & Hongkai Sun
2004 *Language policy and practice in the People's Republic of China: Theory and practice since 1949*. Boston: Springer.

Marie Maegaard, Copenhagen (Denmark)
Karoline Kühl, Copenhagen (Denmark)
Janus Spindler Møller, Copenhagen (Denmark)

41. Surveys

1. Brief history of survey research
2. Survey representation
3. Survey measurement
4. The future of surveys
5. References

1. Brief history of survey research

Surveys are invaluable tools for collecting data on a wide range of topics. Defined as a systematic method for gathering information from a sample of entities for the purpose of constructing quantitative descriptors of a larger population to which those entities belong (Groves et al. 2004), surveys guide many policy decisions.

The earliest form of surveys are government censuses − attempts to count the entire population of a country for political representation, military obligations, or tax purposes. One notable example dating back to 1086 is the *Domesday Book*, a census containing data on every village in England. Censuses continued to play an important role not only administratively, but also ideologically, by classifying populations by race, ethnicity, or language (Kertzer and Arel 2002). Belgium was first to introduce a question on language as a proxy for nationality in its 1846 census (Arel 2001). Prussia and Switzerland followed suit in the 1850s. At the first session of the International Statistical Congress in Brussels, the category 'spoken language' was proposed as part of the standardized census questionnaire, but the recommendation did not become official until the 1972 meeting, where language was proposed as the best objective indicator to capture cultural nationality (Arel 2001).

Surveys became more popular for gathering information on social issues largely because new methods evolved for systematic data collection that provided quicker and cheaper data. In the early 20th century, opinion polls and market research started collecting data via mail and telephone. Compared to face-to-face interviewing, mail questionnaires offered a much cheaper way of measuring literate populations. With the mass penetration of telephone service across the world, researchers saw a new opportunity for a faster way to collect data. By the 1990s, most market and academic research companies in the US had adopted the new mode, and Europe was following suit. An even larger leap in efficiency came from the adoption of computers and the internet, leading to a new mode of data collection: web surveys.

As social surveys were gaining popularity and the modes of data collection expanded, the methods for data collection also changed − posing a fixed set of questions in a standardized manner to respondents selected through probability-based sampling became the norm in the field. The idea of standardizing the questionnaire and the way interviewers asked questions and provided clarifications came about as means of exposing each sampled individual to the same question experience, thus minimizing measurement error. Census-type data collections presented a variety of difficulties related to collecting data from large populations, including analytic ones, and necessitated the use of samples drawn to represent the population of interest. Today, probability samples are an indicator of data quality, and are routinely used by almost all government statistical agencies across the globe.

https://doi.org/10.1515/9783110435351-041

2. Survey representation

2.1. Target population, coverage, and sampling error

The set of persons to be studied is called *target population*. In an ideal world, the listing of all units in the target population is called a *sampling frame*. However, the sampling frame is often not perfectly linked to the target population. For example, in a telephone survey using a frame of telephone numbers, an error arises related to the percentage of the population that cannot be reached by telephone, and how different that part of the population is from the rest on the variables of interest. This type of error is called *coverage error*, and reflects the mismatch between the target population and the existing sampling frame.

Given that a survey cannot possibly be administered to the entire target population, certain techniques are required for drawing a probability sample of the population of interest (assigning each unit a known, nonzero chance of selection), so that inferences to the entire population can be made. The fact that not all members of the target population will be measured subjects the survey estimates to *sampling error*: the possibility that if a new sample is drawn from the same population, it will yield different survey estimates.

Most national and cross-national surveys employ probability samples. The simplest sampling method, *simple random sampling*, assigns an equal probability of selection to each frame element. However, the probabilities of selection do not need to be equal. In fact, various sampling methods exist for overrepresenting specific groups of interest (e.g. ethnic minorities). Several sampling design elements are often considered to minimize the sampling error of a survey statistic: stratification (sorting of the sampling frame into separate subgroups and drawing separate samples from each), clustering (selecting groups of persons rather than individual persons from the frame), and assigning different probabilities of selection to different units in the sampling frame. The Eurobarometer 386 is a good example of a cross-national survey that draws a sample of respondents across 27 member countries to interview them on various emerging issues. The 2012 Special Eurobarometer 386 "Europeans and their Languages" collected responses from about 27,000 respondents as part of a multistage stratified random sample drawn from each country, with probability proportionate to size, representing about 409 million Europeans.

2.2. Nonresponse

Failure to obtain a measurement from the sampled units is referred to as *nonresponse*. Surveys are often more concerned with *unit nonresponse* (the complete failure to obtain a response, i.e. when the person refuses to participate), than with *item nonresponse* (a failure related to the measurement of a specific item). When the nonrespondents are systematically different from the respondents on the statistics of interest, *nonresponse bias* in the survey estimates occurs. Often, it is possible to adjust for such biases using post-survey adjustments, such as weighting and imputation.

3. Survey measurement

3.1. Measurement and survey questions

The main typography of survey questions is by what they measure. Questions that measure an opinion or attitude are referred to as *attitudinal* questions. Since attitudes can change over time and are believed to be strongly dependent on the conversational context (e.g. Tourangeau 1992; Tourangeau, Rips, and Rasinski 2000), there is no objective truth or 'gold standard' against which responses to attitudinal questions can be assessed. In fact, the belief-sampling model (Tourangeau, Rips, and Rasinski 2000) posits that attitudes and opinions are created on the spot, rendering previous responses to the same questions by the same respondents (as in longitudinal data collections) an inadequate gold standard. Thus, measurement of sociolinguistic change (outlined as one of the four goals of sociolinguistic surveys by Lieberson [1980]) may not be best suited through questions collecting opinions and attitudes toward language, as by definition, such opinions and attitudes would be expected to change over time, depending on the provided context.

In contrast, *behavioral* questions, or questions asking what, when, or how often an event has happened, are associated with an objective truth, and when administrative data are available, the accuracy of the reports can be checked against a gold standard. Examples of behavior questions include language use and language acquisition.

Knowledge questions aim to measure respondent's knowledge on the subject of interest, and are often used as screening questions to eliminate respondents who are not knowledgeable about an issue and unlikely to have an opinion. Similar to behavioral questions, there is an objective truth associated with knowledge questions; unlike most behavioral questions (for which administrative data rarely exist), the truth associated with knowledge questions is readily available to researchers. Such questions can be viewed as test questions and used in sociolinguistic surveys to validate self-reported language proficiency.

Finally, *demographic* questions collect information on characteristics such as age, gender, education, race/ethnicity, language spoken at home and at work, etc., and can vary in terms of objectivity/subjectivity. For example, with recent societal developments, demographic characteristics, such as ethnicity or gender, are currently considered subjective, and left to the individual perception of belonging to one group or another, rather than to the presence or absence of certain physical characteristics.

In terms of question structure, survey questions can be divided into four major categories: (1) open-ended (when explicit answer choices are not provided, and respondents use their own words to answer the question); (2) closed-ended (when ordered or unordered response options are presented, and respondents select one or all that apply); (3) partially closed-ended (when closed-ended, unordered response options are provided, followed by 'Other, please specify'); and (4) field coded questions (when the questions are asked as open-ended, and interviewers are responsible for coding the responses to a set of predefined categories). To illustrate, we provide some examples from the Special Eurobarometer 386 "Europeans and their Languages": (1) open-ended ("Thinking about languages other than your mother tongue, which two languages do you think are the most useful for your personal development?"); (2) closed-ended and (3) partially-closed ended ("I am going to read out several ways of learning a foreign language. Please tell

me which of these ways you have ever used: language lessons at school, group language lessons with a teacher, 'one to one' lessons with a teacher, conversation lessons with a native speaker, talking informally to a native speaker …, Other [SPONTANEOUS]"); and (4) field-coded ("Thinking about the languages that you speak, which language is your mother tongue?").

Each type has its advantages and disadvantages. For example, open-ended questions allow respondents to express themselves when all possible responses cannot be anticipated, but in addition to presenting a greater burden for respondents, they require additional investment of time and money to be coded and analyzed. Closed-ended questions are easier to analyze and minimize the cognitive burden for respondents, but also provide opportunities for 'bad' respondent behaviors, such as straight-lining (i.e. selecting the same response category in a series of questions that utilize the same scale), guessing, etc.

3.2. Measurement- and respondent-induced errors

Several theoretical models of the mental processes involved in answering survey questions exist (e.g. Schwarz 1990; Schwarz and Bless 1992; Tourangeau 1984, 1987, 1992). Most models divide response formation into four components: comprehension of the survey question, retrieval of relevant information, use of that information to make a judgment, and selection and reporting of an answer. It is important to note that these components are not necessarily sequential, and that respondents can move back and forth between phases, or even skip one (or more) entirely.

3.2.1. Comprehension

At the comprehension stage, respondents attend to the meaning of words and the question context in order to interpret the question and accompanying instructions. The key issue at this stage is whether the respondents' interpretation of the question matches the researchers' intent. Two intertwined processes are in place at the comprehension stage: understanding of the semantic or literal meaning of the words, and of the pragmatic meaning of the question (Clark and Schober 1992; Strack 1992; Strack and Schwarz 1992). The use of technical or unfamiliar terms, idioms, or ambiguous words can present a comprehension issue. Violations of the assumptions that govern the conduct of everyday conversation (e.g. the cooperative principle, as defined by Grice [1989]) can lead to problems with misinterpretation of the intended meaning(s).

When surveys are administered in different languages, in addition to the direct impact of translation, comprehension problems may occur as a result of differences in cultural norms and practices associated with a particular language. Word meaning and sentence meaning in language comprehension depends on preexisting background knowledge about not only the grammatical norms associated with the language, but also the cultural norms and practices related to it. For example, the question 'Could you close the window?' is not a request for a report on one's physical capability, but a request for an action which reflects a norm of politeness.

Languages differ in their context dependency, and this is reflected in conversational norms across cultures; for example, many words in Chinese acquire meaning only in the

conversational context and cannot be translated directly. Such context dependency can have consequences for question interpretation when partially redundant information is presented. For example, Haberstroh et al. (2002) asked Chinese and German students about their academic and general life satisfaction, flipping the order of the questions. As expected, in contrast to German students, who used information brought to mind by the academic satisfaction question to answer the general life satisfaction question, Chinese students, being more sensitive to the conversational context and detecting potentially redundant questions, disregarded the information on academic life when reporting on general life satisfaction.

Languages also have different inherent frames of reference for describing relationships among objects. A common distinction in psycholinguistics is between egocentric and allocentric languages. Such intrinsic differences may affect interpretation of survey visual images or response scales. For example, a vertical or horizontal scale orientation may influence whether response categories are perceived across languages as being similar or distinct.

3.2.2. Retrieval

At the retrieval stage, respondents recall information relevant to answering the question. Such information is often retrieved from long-term memory as well as from general knowledge. For behavioral questions, relevant information is likely to be recalled if the event is of personal importance and is rather unique (i.e. does not occur frequently); see Schwarz (1990) and Bradburn, Rips, and Shevell (1987). Aside from the nature of the events in question (and often, the respondent's age), recall and its accuracy are also affected by how long ago events took place (Cannell, Miller, and Oksenberg 1981; Loftus et al. 1992; Means et al. 1989; Smith and Jobe 1994), the time on task (Williams and Hollan 1981), and the number of cues presented to help trigger the retrieval process (Tulving and Thompson 1973). Due to a phenomenon known as *language-dependent recall*, language has been demonstrated to serve as a powerful retrieval cue in bilingual respondents: a match between language of encoding (the language in which information was committed to long term memory) and language of recall (the language in which the survey question is asked) ensures not only more memories (Schrauf and Rubin 1998), but also induces a specific mind-set, resembling the process assumed to underlie state-dependent memory (Marian and Neisser 2000).

Recall for attitudinal questions depends on whether an opinion has been formed before and is accessible at the time of the interview. The personal importance of the issue, the recency of the formed opinion, and whether respondents have been asked before a related question are key factors which determine the likelihood of recall (Sudman, Bradburn, and Schwarz 1996). Language of survey administration can also determine what information is most accessible at any given time by activating the cognitive-affective cultural framework associated with it (Schrauf 2000). For example, a study of Greek students attending an American school in Greece showed that the correlation between responses to the same attitudinal questions administered in English and in Greek was low for domains in which American and Greek norms differed in what was considered socially desirable, and high for domains where cultural values converged (Triandis et al. 1965). Similar findings have been reported by Marin et al. (1983).

Language codability (i.e. the ease with which a language can express a concept) has also been found to affect cognition and how people answer survey questions. Not surprisingly, the most highly codable concepts are represented by the most frequently used words, which are short, and easy to write and pronounce (see Whitney 1998). Codability has been demonstrated to affect retrieval (Lucy 1992; Lucy and Shweder 1979; Lucy and Wertsch 1987) and comparative judgment (Kay and Kempton 1984).

3.2.3. Judgment

At the judgment stage, respondents assess the retrieved information and fill in the gaps. In the case of attitudinal questions, if a preexisting opinion is not recalled, judgments have to be computed taking into account all recalled relevant information. If an opinion has been recalled, it may not exactly match the question intent, and thus may need to be recomputed. Similarly, for most behavioral questions, respondents may need to adjust the recalled information to be fitted to the question; for example, respondents need to determine whether the recalled events fit the question reference period, or calculate the number of recalled events. Most respondents, however, are 'cognitive misers' (Taylor 1981), and do not retrieve all potentially relevant information. Instead, they stop searching as soon as they have just enough information, relying on the information that comes to mind most easily, such as information that is temporarily accessible in memory due to the preceding questions. In addition to such context dependency, the accessibility of relevant information depends on the frequency of its use − when respondents think about an issue frequently, related information becomes chronically accessible in their memories.

Language codability affects judgement; for example, the existence of a stereotype in a language influences one's impressions about a person (Hoffman, Lau, and Johnson 1986). In addition, language's ability to cue the culture associated with it would determine whether a survey question is perceived as sensitive or not (Marin et al. 1983; Triandis et al. 1965).

3.2.4. Response formatting and editing

At the response formatting stage, the respondent selects and communicates a response. This involves mapping the retrieved information/formed judgment to the response options provided by the question, often formatting the response. This formatting sometimes means altering the response in order to provide a socially acceptable answer, or to be consistent with previous responses. For example, instead of reporting the actual number of alcoholic beverages consumed per week, respondents may intentionally modify their true number in order to comply with cultural values and norms in place. Such response editing is more pronounced in interviewer as opposed to in self-administered surveys (for a review, see DeMaio 1984). Furthermore, language of survey administration can prime a cultural identity that would determine how a survey question is perceived, and what constitutes a 'socially appropriate' response.

The response options to which respondents are trying to map their judgment often convey unintended information. For example, the middle of a rating scale may be perceived as the population average (Schwarz and Hippler 1991). Furthermore, depending

on the mode of data collection (visual or auditory), respondents may endorse response options presented first on the list (primacy effect), or the ones presented last on the list (recency effect). Such response order effects have been reported for both attitudinal and behavioral questions (Schwarz, Hippler, and Noelle-Neumann 1992, 1994).

3.3. Measurement and mode of data collection

Traditionally, surveys relied on three basic modes of data collection: face-to-face interviewing, telephone interviewing, and mail administration. The computer revolution added nuances to these traditional methods (e.g. Computer Assisted Personal Interviewing, or CAPI, and Computer Assisted Telephone Interviewing, or CATI), as well as introduced new modes of data collection, such as web surveys, in which a computer administers the survey questions on a website.

The main distinction across modes of data collection is the level of interviewer involvement, which has many implications for data accuracy and cost. Interviewers can be effective in recruiting sample members, which minimizes nonresponse error. However, their social presence and demographic characteristics can have undesirable effects on measurement, especially as they relate to sensitive or socially desirable questions. A considerable body of experimental research exists on the differences between face-to-face and mediated interactions. Research on nonverbal communications suggests that nonverbal messages communicated by facial expressions and gestures add a nuance to communication beyond spoken language (e.g. Ekman 1965). Another important mode difference between face-to-face and telephone surveys is the pace of the interview. Studies comparing how long it takes to complete a series of questions in both modes report shorter times for telephone surveys, likely related to the norms of telephone conversation and respondents' reluctance to engage in a lengthy discussion with a stranger on the phone (Groves 1978; Kormendi 1988; Sykes and Collins 1988). From a data quality perspective, this translates into reduced detail in answers to open-ended questions (especially from highly educated respondents [Groves 1978; Kormendi 1988]), increased tendency to agree and/or to select response options on the extreme end of the scale, and increased item-missingness (Jordan, Marcus, and Reeder 1980).

3.4. Measurement and interviewer influences

Interviewers can have unexpected effects on data accuracy. Such influences can occur in two main ways: (1) through interviewers' sociodemographic characteristics (both observable and unobservable) and (2) through interviewers' individual behaviors (e.g. skipping or rewording a question, incorrectly recording a response, emphasizing certain parts of questions more than others, and/or providing clarifications). Observable characteristics, such as race and gender, have been demonstrated to have an effect on survey responses when questions that pertain to such demographic characteristics are asked (e.g. Groves and Fultz 1985; Schaeffer 1980). Unobservable characteristics, such as interviewer experience, can also have an effect on respondents' answers; for example, more experienced interviewers have been reported to yield lower rates of item-missingness on sensitive questions (Stevens and Bailar 1976). Interviewers' behaviors associated with

failure to read the question exactly as written, inability to follow skip logic, or reading a question too fast have been shown to contribute to errors in the collected data.

One finding in the survey literature is that respondents interviewed by the same interviewer tend to provide more similar responses; this has been explained by the fact that respondents react to these observable and unobservable characteristics, often unintentionally. This is particularly true for sensitive or attitudinal questions (e.g. Bailar, Bailey, and Stevens 1977; Fellegi 1964; Schnell 1997), but interestingly, the effects have also been found for factual items (such as age and ethnicity) and for self-administered items administered while the interviewer is in the room (Groves and Magilavy 1986; Mangione, Fowler, and Louis 1992; O'Muircheartaigh and Campanelli 1998; West and Peytcheva 2014).

3.5. Measurement and language influences

There are two main mechanisms through which the language of survey interviews may influence survey responses: cultural frame switching and language-dependent recall. Cultural frame switching refers to a phenomenon where individuals with more than one cultural identity (typically, bilinguals), move between cultural meaning systems depending on situational cues. Each cultural meaning system serves as an interpretive frame that affects individual's cognition, emotion, and behavior (Geertz 1993; Kashima 2000; Mendoza-Denton et al. 1999). As discussed earlier, language can serve as a situational cue for the cultural system associated with it, which may prompt bicultural and/or bilingual respondents to different question interpretations or different judgments based on the cultural frame induced by a given language (e.g. Marin et al. 1983; Ross, Xun, and Wilson 2002; Trafimow et al. 1997).

Similarly, language-dependent recall has been demonstrated in several bilingual groups in terms of number of recalled events, time in life when the events took place, and language associated with the events, showing that memories encoded in a particular language were easier to recall in the same language (e.g. Bugelski 1977; Marian and Neisser 2000; Schrauf and Rubin 1998; Schrauf and Rubin 2000). Such findings suggest that the language of survey administration would affect both the quality and quantity of recall, and surveys employing more than one language (where the expected proportion of respondents speaking more than one language is high), may need to rethink the current procedure of asking questions in all domains in the same language.

4. The future of surveys

Surveys currently face many challenges, but probably the most worrisome is the increasing reluctance of sample persons to participate, resulting in a potential threat of nonresponse bias in survey estimates. Growing nonresponse rates challenge researchers' control over data collection costs and optimal design prior to initiation of data collection. Responsive survey design addresses these problems by altering the survey design or data collection procedures *during* data collection in order to maintain cost efficiency and reduce survey error (Groves and Heeringa 2006). Responsive designs are dependent on

the design features of a survey and data collection environment; therefore, they are study-specific. Such designs have gained popularity over the past decade, and may become the new norm for designing survey data collections.

Declining response rates, increasing costs, coverage issues, and data collection deadlines have also led to increased use of mixed-mode survey designs. The use of a combination of data collection methods reduces cost, as it typically involves an attempt to collect data in a cheaper mode (e.g. mail), followed by a more expensive mode (e.g. telephone), and possibly moving to an even more costly mode (e.g. face-to-face interviewing). The American Community Survey is an example of this approach: it starts in mail mode, then telephone follow-up with nonrespondents, and finally, face-to-face follow-up with a subsample of remaining nonrespondents (Alexander and Wetrogan 2000). The attractiveness of mixed-mode designs lies in their ability to mitigate the weaknesses of individual modes. For example, to reduce coverage bias in the early days of telephone data collection, mixed-mode dual frame designs were often employed, benefiting from the cost savings of telephone interviewing and the complete coverage of face-to-face data collection (for a detailed discussion, see Groves and Lepkowski 1985). Another feature that makes mixed-mode designs attractive is their application in reducing nonresponse bias. Since nonresponse includes both respondents not contacted and those who refused to cooperate under the initial protocol, implementing a different mode of data collection can be addressed both by changing the method of contact and using different persuasive techniques, particularly through the use of interviewers. It is not necessarily true that some modes are better than others for a particular population. To the extent that individuals vary in their likelihood of participation across modes, and respondents to different modes are somewhat different, the threat of nonresponse bias is minimized through the use of multiple modes.

Finally, the amount of information currently available from different data sources can be linked and analyzed to supplement any survey data collection. The concept of 'big data' has recently gained popularity, representing the collection, techniques, manipulation, and analyses of enormous amounts of data. The utility of big data for gaining insights into social behavior and attitudes on its own is poorly understood. Recent studies have compared estimates from big data sources to estimates from traditional surveys; however, there is still not enough research to inform best practices for using big data to produce population estimates. One of the main criticisms against big data is the lack of theory for drawing inferences − data are often generated as a byproduct from various processes, and often there is not sufficient information about that. A useful strategy might be to combine, rather than replace, traditional survey data with new data sources.

5. References

Alexander, Charles H., Jr & Signe Wetrogan
2000 Integrating the American community survey and the intercensal demographic estimates program. https://census.gov/content/dam/Census/library/working-papers/2000/acs/2000_Alexander_01.pdf/. (accessed on 14 September 2018).

Arel, Dominique
2001 Language categories in censuses: Backward- or forward-looking? In David Kertzer & Dominique Arel (eds.), *Census and identity: The politics of race, ethnicity, and language in national censuses*, 92−120. Cambridge: Cambridge University Press.

Bailar, Barbara, Leroy Bailey & Joyce Stevens
1977 Measures of interviewer bias and variance. *Journal of Marketing Research* 14(3). 337–343.
Bradburn, Norman, Lance Rips & Steven Shevell
1987 Answering autobiographical questions: The impact of memory and inference on surveys. *Science* 236(4798). 157–161.
Bugelski, Richard
1977 Imagery and verbal behavior. *Journal of Mental Imagery* 1. 39–52.
Cannell, Charles, Peter Miller & Lois Oksenberg
1981 Research on interviewing techniques. *Sociological methodology* 12. 389–437.
Clark, Herbert & Michael Schober
1992 Asking questions and influencing answers. In Judith Tanur (ed.), *Questions about questions: Inquiries into the cognitive basis of surveys*, 15–48. New York: Russell Sage Foundation.
DeMaio, Theresa
1984 Social desirability and survey measurement: A review. In Charles Turner & Elizabeth Martin (eds.), *Surveying subjective phenomena*, vol. 2, 257–281. New York: Russell Sage Foundation.
Ekman, Paul
1965 Differential communication of affect by head and body cues. *Journal of Personality and Social Psychology* 2(5). 726–35.
Eurobarometer 386
2015/2016 Special Eurobarometer 386: Europeans and their languages. http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_386_en.pdf/. (accessed on 14 September 2018).
Fellegi, Ivan
1964 Response variance and its estimation. *Journal of the American Statistical Association* 59. 1016–1041.
Geertz, Clifford
1993 *The interpretation of cultures*. New York: Basic Books.
Grice, Paul
1989 *Studies in the way of words.* Cambridge MA: Harvard University Press.
Groves, Robert
1978 On the mode of administering a questionnaire and responses to open-ended items. *Social Science Research* 7(3). 257–271.
Groves, Robert & James Lepkowski
1985 Dual frame, mixed-mode survey designs. *Journal of Official Statistics* 1(3). 263–286.
Groves, Robert & Lou Magilavy
1986 Measuring and explaining interviewer effects in centralized telephone surveys. *Public Opinion Quarterly* 50(2). 251–266.
Groves, Robert & Nancy Fultz
1985 Gender effects among telephone interviewers in a survey of economic attitudes. *Sociological Methods and Research* 14(1). 31–52.
Groves, Robert & Steven Heeringa
2006 Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 169(3). 439–457.
Groves, Robert, Floyd Fowler, Mick Couper, James Lepkowski, Eleanor Singer & Roger Tourangeau
2004 *Survey methodology.* Hoboken, NJ: Wiley & Sons.
Haberstroh, Susanne, Daphna Oyserman, Norbert Schwarz, Ulrich Kuhnen & Li-Jun Ji
2002 Is the interdependent self more sensitive to question context than the independent self? Self-construal and the observation of conversational norms. *Journal of Experimental Social Psychology* 38(3). 323–329.

Hoffman, Curt, Ivy Lau & David Johnson
1986 The linguistic relativity of person cognition: An English-Chinese comparison. *Journal of Personality and Social Psychology* 51. 1097–1105.
Jordan, Lawrence, Alfred Marcus & Leo Reeder
1980 Response styles in telephone and household interviewing: A field experiment. *Public Opinion Quarterly* 44(2). 210–222.
Kashima, Yoshihisa
2000 Conceptions of culture and person for psychology. *Journal of Cross-Cultural Psychology* 31(1). 14–32.
Kay, Paul & Willett Kempton
1984 What is the Sapir-Whorf hypothesis? *American Anthropologist* 86. 65–79.
Kertzer, David & Dominique Arel
2001 Censuses, identity formation, and the struggle for political power. In David Kertzer & Dominique Arel (eds.), *Census and identity: The politics of race, ethnicity, and language in national censuses*, 1–42. Cambridge: Cambridge University Press.
Kormendi, Eszter
1988 The quality of income information in telephone and face-to-face surveys. In Robert Groves, Paul Biemer, Lars Lyberg, James Massey, William Nichols & Joseph Waksberg (eds.), *Telephone survey methodology*, 341–356. New York: Wiley & Sons.
Lieberson, Stanley
1980 Procedures for improving sociolinguistic surveys of language maintenance and language shift. *International Journal of the Sociology of Language* 25. 11–27.
Loftus, Elizabeth, Kyle Smith, Mark Klinger & Judith Fiedler
1992 Memory and mismemory for health events. In Judith Tanur (ed.), *Questions about questions: Inquiries into the cognitive basis of surveys*, 102–137. New York: Russell Sage Foundation.
Lucy, John
1992 *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
Lucy, John & Richard Shweder
1979 Whorf and his critics: Linguistic and nonlinguistic influences on color memory. *American Anthropologist* 81. 581–615.
Lucy, John & James Wertsch (eds.)
1987 Vygotsky and Whorf: A comparative analysis. In Maya Hickmann (ed.), *Social and functional approaches to language and thought*, 67–86. Orlando, FL: Academic Press.
Mangione, Thomas, Floyd Fowler & Thomas A. Louis
1992 Question characteristics and interviewer effects. *Journal of Official Statistics* 8(3). 293–307.
Marian, Viorica & Ulric Neisser
2000 Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General* 129(3). 361–368.
Marin, Gerardo, Harry Triandis, Hector Betancourt & Yoshihisa Kashima
1983 Ethnic affirmation versus social desirability: Explaining discrepancies in bilinguals' responses to a questionnaire. *Journal of Cross-Cultural Psychology* 14(2). 173–186.
Means, Barbara, Arti Nigam, Marlene Zarrow, Elizabeth Loftus & Molla Donaldson
1989 *Autobiographical memory for health-related events: Vital and health statistics*. Washington, D.C.: US Government Printing Office.
Mendoza-Denton, Rodolfo, Yuichi Shoda, Ozlem Ayduk & Walter Mischel
1999 Applying cognitive-affective processing system (Caps) theory to cultural differences in social behavior. In Walter J. Lonner & Dale L. Dinnel (eds.), *Merging past, present, and future in cross-cultural psychology: Selected papers of the 14th international congress*

of the International Association for Cross-Cultural Psychology, 205–217. Lisse, The Netherlands: Swets & Zeitlinger.

O'Muircheartaigh, Colm & Pamela Campanelli
1998 The relative impact of interviewer effects and sample design effects on survey precision. *Journal of the Royal Statistical Society, Series A* 161(1). 63–77.

Ross, Michael, Elaine Xun & Anne Wilson
2002 Language and the bicultural self. *Personality and Social Psychology Bulletin* 28(8). 1040–1050.

Schaeffer, Nora-Cate
1980 Evaluating race-of-interviewer effects in a national survey. *Sociological Methods and Research* 8(4). 400–419.

Schnell, Rainer
1997 *Nonresponse in Bevölkerungsumfragen*. Opladen: Leske + Budrich.

Schrauf, Robert
2000 Bilingual autobiographical memory: Experimental studies and clinical cases. *Culture and Psychology* 6(4). 387–417.

Schrauf, Robert & David Rubin
1998 Bilingual authobiographical memory in older adult immigrants: A test of cognitive explanations of the reminiscence bump and the linguistic encoding of memories. *Journal of Memory and Language* 39(3). 437–457.

Schrauf, Robert & David Rubin
2000 Internal languages of retrieval: The bilingual encoding of memories for the personal past. *Memory and Cognition* 28(4). 616–623.

Schwarz, Norbert
1990 Assessing frequency reports of mundane behaviors: Contributions of cognitive psychology to questionnaire construction. In Clyde Hendrick & Margaret Clark (eds.), *Research methods in personality and social psychology*, 98–119. Newbury Park, CA: Sage.

Schwarz, Norbert & Hans Hippler
1991 Response alternatives: The impact of their choices and ordering. In Paul Biemer, Robert Groves, Nancy Mathiowetz & Seymor Sudman (eds.), *Measurement errors in surveys*, 41–56. Chichester: Wiley.

Schwarz, Norbert & Herbert Bless
1992 Constructing reality and its alternatives: Assimilation and contrast effects in social judgment. In Leonard Martin & Abraham Tesser (eds.), *The construction of social judgment*, 217–245. Hillside, NJ: Erlbaum.

Schwarz, Norbert, Hans Hippler & Elisabeth Noelle-Neumann
1992 A cognitive model of response order effects in survey measurement. In Norbert Schwarz & Seymor Sudman (eds.), *Context effects in social and psychological research*, 187–201. New York: Springer-Verlag.

Schwarz, Norbert, Hans Hippler & Elisabeth Noelle-Neumann
1994 Retrospective reports: The impact of response alternatives. In Norbert Schwarz & Seymor Sudman (eds.), *Autobiographical memory and the validity of retrospective reports*, 187–202. New York: Springer.

Smith, Albert & Jared Jobe
1994 Validity of reports of long-term dietary memories: Data and a model. In Norbert Schwarz & Seymor Sudman (eds.), *Autobiographical memory and the validity of retrospective reports*, 121–140. New York: Springer.

Stevens, Joyce & Barbara Bailar
1976 The relationship between various interviewer characteristics and the collection of income data. In *Proceedings of the social statistics section: American Statistical Association*, 785–790.

Strack, Fritz
1992 Order effects in survey research: Activation and informative functions of preceding questions. In Norbert Schwarz & Seymor Sudman (eds.), *Context effects in social and psychological research*, 23–34. New York: Springer.
Strack, Fritz & Norbert Schwarz
1992 Implicit cooperation: The case of standardized questioning. In Gun Semin & Klaus Fiedler (eds.), *Language, interaction, and social cognition*, 173–193. Newbury Park, CA: Sage.
Sudman, Seymor, Norman Bradburn & Norbert Schwarz
1996 *Thinking about answers*. San Francisco, CA: Jossey-Bass Publishers.
Sykes, Wendy & Martin Collins
1988 Effects of mode of interview: Experiments in the UK In Robert Groves, Paul Biemer, Lars Lyberg, James Massey, William Nichols & Joseph Waksberg (eds.), *Telephone survey methodology*, 301–320. New York: Wiley & Sons.
Taylor, Shelley
1981 The interface of cognitive and social psychology. In John Harvey (ed.), *Cognition, social behavior, and the environment*, 189–211. Hillsdale, NJ: Erlbaum.
Tourangeau, Roger
1984 Cognitive science and survey methods: A cognitive perspective. In Thomas Jabine, Miron Straf, Judith Tanur & Roger Tourangeau (eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines*, 73–100. Washington, DC: National Academy Press.
Tourangeau, Roger
1987 Attitude measurement: A cognitive perspective. In Hans Hippler, Norbert Schwarz & Seymor Sudman (eds.), *Social information processing and survey methodology*, 149–162. New York: Springer.
Tourangeau, Roger
1992 Attitudes as memory structures: Belief sampling and context effects. In Norbert Schwarz & Seymor Sudman (eds.), *Context effects in social and psychological research*, 35–47. New York: Springer.
Tourangeau, Roger, Lance Rips & Keneth Rasinski
2000 *The psychology of survey response*. New York: Cambridge University Press.
Trafimow, David, Ellen Silverman, Ruth Mei-Tai Fan & Josephine Law
1997 The effect of language and priming on the relative accessibility of the private self and the collective self. *Journal of Cross-Cultural Psychology* 28(1). 107–123.
Triandis, Harry, Earl Davis, Vasso Vassiliou & Maria Nassiakou
1965 *Some methodological problems concerning research on negotiations between monolinguals* (Department of Psychology Technical report #28). Urbana: University of Illinois.
Tulving, Endel & Donald Thompson
1973 Encoding specificity and retrieval processes in episodic memory. *Psychological Review* 80(5). 352–373.
West, Brady & Emilia Peytcheva
2014 Can interviewer behaviors during ACASI affect data quality? *Survey Practice* 7(5). 1–10.
Williams, Michael David & James D. Hollan
1981 The process of retrieval from very long-term memory. *Cognitive Science* 5(2). 87–119.
Whitney, Paul
1998 *The psychology of language*. Boston, MA: Houghton Mifflin.

Emilia Peytcheva, Chapel Hill (USA)

42. The representativeness of samples

1. Introduction
2. A historical sketch of the concept of 'representativeness' in statistics
3. The definition of 'representativeness of samples'
4. Errors that distort the representativeness of samples
5. Not all samples have to be representative
6. References

1. Introduction

According to Wolfram (2011: 296), the term "'sociolinguistic fieldwork' has been assigned to cover a wide variety of activities and procedures for collecting speech data, extending from the incidental observation of language use in the course of everyday interaction to the systematic elicitation of structural responses conducted within an experimental study". Insofar as sociolinguistic research is based on quantitative analyses of surveys, the *representativeness* of results is important. Early empirical research in this field focusing on historical reconstruction of previous language forms was mainly based on observations among purposefully selected older, male, uneducated, lifetime residents (Milroy and Gordon 2003: 12). Selection techniques for sample participants had to be adapted when, during the 20th century, the focus of research shifted toward description of contemporary language. On the one hand, due to the homogeneity of certain interesting language features, even a single purposefully observed speaker has, or a few consciously selected groups of residents have, provided representative information about that specific feature; on the other hand, for research on more general language phenomena, probability samples have set the standard for valid empirical results since Labov's classic sociolinguistic studies in the 1960s (Labov 1984; Wolfram 2011: 299). Hence, similar to other fields of empirical studies, the sampling design to be applied must be tailored to the research questions of interest.

From the point of view of applied statistics, and with respect to the quality of the survey results, the question of their representativeness is of the utmost importance. The purpose of this chapter is to provide readers with some general information regarding this topic. It includes a number of observations that are also of relevance to (young) researchers who would like to investigate language contact phenomena by making use of surveys.

The following section provides a historical sketch of the use and meaning of the term 'representativeness' in the field of statistics. Section 3 presents a practical and helpful definition of the term that can serve as a label of quality for conclusions drawn from a sample to a population. Furthermore, various sampling strategies are considered from this point of view. Section 4 discusses various sources of non-sampling errors which may distort the representativeness of a survey, and some methods to repair such distortion are considered. In the concluding section, it is noted that the empirical results do not always have to be representative, giving rise to the term 'informative samples'.

https://doi.org/10.1515/9783110435351-042

2. A historical sketch of the concept of 'representativeness' in statistics

The use of the term 'representativeness' regarding the quality of statistical surveys has a long tradition, beginning at the end of the 19th century. During that century, the results in the field of official statistics were based almost exclusively on full population surveys, or censuses. In light of the extraordinarily high expenditure for censuses, Anders N. Kiaer, who was the first director of the Norwegian Statistical Office (serving from 1867 to 1913), and responsible as such for the decennial censuses in Norway, organized complementary partial investigations of the Norwegian population that were carried out for the first time in 1894 (see here, and in Bellhouse [1988], Bethlehem [2009], and Kruskal and Mosteller [1980]). In the following year, Kiaer presented his concept of the 'representative method' at the congress of the International Statistical Institute (ISI) in Berne, Switzerland. The idea of the representative sample was that it was an approximate miniature of the population with respect to various characteristics from known census results. Kiaer stressed that the main advantage of his approach was that a representative sample could provide more information on more variables than a census could. Although his idea has a rather intuitive and non-probabilistic approach, its influence cannot be underestimated, because it marks not only the beginning of the use of the term 'representativeness' in the field of sampling, but also of sampling theory itself.

In 1901, the statistician and economist Ladislaus von Bortkiewicz formulated a proposal to check the representativeness of a sample by analyzing certain population characteristics, and quantifying their agreement with the respective population parameters. Subsequently, Lucien March discussed the need for randomness in sample selection regarding the representativeness of estimates. In particular, he described simple random sampling of individual units and of clusters of such units. A prerequisite for the application of the representative method is that the selection process should be fully described in the publication of the sample results. This claim has since lost none of its significance, but remains widely ignored in scientific publications. In 1906, the British mathematician Arthur L. Bowley pointed out that probability theory had developed rapidly in the preceding centuries, yet had not been applied to several practical problems to which it might be of use; he mentioned the method of sampling among these topics (Bellhouse 1988: 5). Whereas Kiaer had only intuition, Bowley argued his issues from a theoretically grounded perspective. Eventually, at the ISI meeting of 1925, a resolution was passed which accepted as practically applicable both random sampling with equal inclusion probabilities and purposive sampling in cases where representative sample results are required.

An outstanding contribution to the field of sampling theory regarding representativeness of samples was a paper by Jerzy Neyman (1934). This work presented a theory for stratified random sampling with optimal allocation of the total sample size on given population strata, which left the common path of equal probability random sampling to achieve the representativeness of results by the representativeness of the entire sample. The sample no longer had to be a strict miniature of the entire population for all variables, but rather had to represent the study characteristics in an effective way.

Although the term 'representativeness' has since then been widely used in the presentation of survey results (one need only read newspapers to observe this), seemingly there is still a need to define its concrete meaning in order for it to be usefully applied. This

has not changed much since the end of the 1970s, when Kruskal and Mosteller (1980: 169–172) classified the meaning of the term as used in different publications, resulting in the following nine (only partially overlapping) categories:

1. 'Representativeness' is sometimes used only because of its suggestive effect which underscores that the reader should have no doubt about the 'high quality' and 'reliability' of a study.
2. 'Representativeness' is used to mean the absence of selective forces that might have an effect on the generalizability of the survey results.
3. Another meaning of the 'representativeness' of a sample is that the sample is regarded as a miniature of the population in terms of all (or at least some) important characteristics.
4. A different application example of the term 'representativeness' is its use as a synonym for the purposive selection of elements typical of the entire population regarding the characteristics of interest.
5. The term 'representativeness' is sometimes used in the sense that the heterogeneity of the population units is covered by the sample.
6. Furthermore, 'representativeness' is used as a vague term in the summary of a study to avoid the description of the entire selection and estimation process.
7. A 'representativeness' statement has also been used as a description for samples drawn randomly, with equal inclusion probabilities for all population units.
8. Another use of the term 'representativeness' is to describe a survey in which reliable estimates of the population characteristics of interest are permitted.
9. Finally, the 'representativeness' of a sample may simply mean that the sample is good enough for its specific purpose, which may not necessarily require an accurate estimation of a given population parameter.

The subsequent section presents a practically applicable definition that may serve as an indicator of a high-quality sample survey. This definition was developed from studies by Gabler (1996), Quatember (1996), and Gabler and Quatember (2012).

3. The definition of 'representativeness of samples'

A statistical sample survey is carried out with the aim of obtaining information from a finite set of units (called the population) regarding the (uni- or multidimensional) frequency distributions of study variables or parameters characterizing these population distributions. Examples of such populations may be residents of a certain region, pupils of a particular birth year in various OECD countries, or eligible voters in a federal state. Some variables of interest may be the characteristics of language use in a region, the scores of students on competency tests (such as the PISA survey), or the voting behavior of voters; some examples of parameters of interest may be the number of people speaking a certain dialect or possessing another language attribute, the distribution and average of students' performance or scores on exams, the current proportions of different political parties in opinion polls, or any other statistical indicator characterizing the relevant populations with respect to the topics of interest.

Efficient conclusions based on calculated sample estimates of unknown, yet of interest, population characteristics requires the application of an adequate sampling design

which consists of both the techniques for survey participant selection and the methods for estimating the parameters of interest on the basis of the observed data. In this sense, the sample (and/or the survey) need not be representative as a whole, but must be representative of the population in terms of the population characteristics of interest.

Hence, given a well-designed definition, the term 'representativeness of a sample' should serve as an indicator of the good quality of statistical sample survey outcomes. For this reason, a sample is described as: 'exactly representative' with respect to a certain population distribution or parameter, if this characteristic is exactly reproducible from the sample; 'representative' with respect to a certain population distribution or parameter, if this characteristic can be estimated (at least approximately) from the observed sample data in an unbiased way, and the estimation satisfies a previously defined accuracy requirement; and 'not representative' in this respect at all, if it is neither exactly representative nor representative regarding the population characteristic.

According to this definition, the term 'exactly representative' corresponds to the optimum target when it is applied to the distribution or parameter of interest. Of course, in the practice, such a perfect population miniature with respect to the questions of interest usually cannot be achieved in a sample except by chance. However, this term can be used to describe certain aspects of the applied sampling design. For example, the population distribution of a certain auxiliary variable is exactly represented in stratified simple random sampling with proportional allocation of the total sample size to predefined strata, in the purposive selection scheme of quota sampling, or in application of estimation methods (such as ratio or regression estimation) for which known population parameters are used. Such samples are also described as 'balanced' in terms of these characteristics (Deville and Tillé 2004: 895).

In the given definition, the representativeness of a sample regarding a distribution or parameter is based on the theoretical concept of the statistical unbiasedness of estimators (Särndal, Swensson, and Wretman 1992: 40) and on the efficiency of estimates. The postulated approximate unbiasedness of an estimator means we can be sure that, on average, the estimate equals (at least approximately) the parameter under study. To additionally achieve a predefined accuracy under the given sampling design, a sufficiently large sample size is required. For specific types of statistical analyses, several 'rules of thumb' exist in this respect. For the estimation of parameters, such as proportions or means, the necessary sample sizes can be calculated, given a desired range of statistical fluctuation once simple probability sampling schemes are applied (Lohr 2010: 46–50).

According to this definition, the representativeness of a sample can serve as an indicator of high-quality inference from sample to population on a probability basis, which implicitly requires the use of an adequate sampling technique. Clearly, this demand is satisfied only by probability sampling methods, such as simple, stratified, cluster, two-stage, or probability proportional to size random sampling, with their known nonzero sample inclusion probabilities for all population units, and under the ideal conditions of the so-called 'urn model'. These probabilities are needed to allow calculation of the so-called 'design weights' for the strictly design-based estimation process (Quatember 2015: 11–23).

In a simple random sample with equal inclusion probabilities for all population units, for instance, when a large enough sample size is chosen, the design-based estimation of the proportion of people possessing a certain language attribute by the sample proportion satisfies the requirements for representativeness regarding this parameter. The same applies to the proportional stratified simple random sampling scheme, which is exactly

representative for the stratification variable and representative for all other variables at a lower sample size when the stratum and study variable are associated. Therefore, it is important to use stratification variables that are related to the study questions regarding, for instance, language behavior. For this purpose, sex, age, class, education, socioeconomic status, and ethnicity have been used as nonlinguistic variables (Milroy and Gordon 2003: 18). In an empirical study of language interactions, for instance, compared with simple random sampling of population units, random cluster, or two-stage, sampling of blocks of households usually achieves representativeness regarding parameters as a proportion of a given attribute only at larger sample sizes and lower field costs (when the clusters differ in the distribution of the study variable). In unequal probability sampling schemes, such as probability proportional to size sampling, the unequal inclusion probabilities of different selection units are compensated by the units' design weights in the estimation process, resulting in estimates that are all the more accurate, the more the size of the variable is related to the study variable (such as tax yield).

These inclusion probabilities, which make possible projection from sample data to population characteristics, certainly cannot be calculated for non-probability sampling procedures, such as quota, snowball, or expert choice sampling, or for non-probability-based self-selection online surveys, in which the internet is used as the selection procedure itself, and not as a survey mode for probability-based sampling procedures (Quatember 2015: 23–25). For instance, in a quota sample, which is increasingly used for empirical research in different fields, and in which the distribution of the chosen quota variables, such as sex, age, and education level, is exactly represented, a strictly design-based estimator (such as in a proportional, stratified random sample) cannot be calculated, because the inclusion probabilities necessary for such a calculation are unknown. Calculating an estimate of a parameter of interest in a quota sample based on the formulas of a probability sampling method (such as simple or proportional stratified random sampling) simply means that the probability sampling method is assumed to be a model for the selection process of the non-probability quota technique. The question regarding representativeness of the survey results then becomes how the real selection process actually is fitted to the presumed selection model (Quatember 2019).

Expert choice sampling has consistently been used in observations of typical groups in empirical studies (Labov 1984: 48–49). Furthermore, network analyses are based on methods like snowball sampling, in which the researcher identifies a few people as sample units of the population of interest, who then identify other possible study units (Dörnyei 2007: 98). Again, if representative pictures of certain population characteristics are to be painted, a realistic model of the true selection process must be assumed. According to our definition of the term, the representativeness of survey results must be rejected when this model does not hold.

There is also a high risk of nonrepresentative results when opt-in online surveys based on self-selection of survey units are carried out, because selection bias can hardly be avoided, except for populations defined as the web survey participants themselves. Dörnyei (2007: 121) lists several benefits of web surveys, including reduced cost, ease of administration, respondent anonymity, and international access. At the same time, their limitations regarding the representativeness of survey results must be considered when, for instance, the interviewees are simply recruited among users who visit the web page with the questionnaire merely by chance. The same applies when looking at language data provided by Facebook or Google users. Although one may be blinded by

the sheer mass of available information when extrapolating conclusions based on these observations to a target population, the problem of modeling the design weights of the population units must be considered. The projection quality of such samples might potentially be improved by 'calibrating' them to known population distributions of some (sociodemographic) variables associated with the research questions. This can be done by applying methods such as post-stratification (Lohr 2010: 342–345).

Consequently, the representativeness of non-probability samples regarding the research questions of interest depends on the validity of (oftentimes more than) doubtful models. Such representativeness was apparently not achieved when the American weekly magazine *Literary Digest* asked its own subscribers and persons from existing lists of automobile and telephone owners to participate in an opinion poll right before the US presidential elections of 1936 (Squire 1988). Over two million voters took part in the survey. The estimates of the election results, based on their 'votes', falsely predicted a clear victory for the Republican candidate. This illustrates that the size of the sample alone does not at all guarantee quality when the selection of respondents itself (and the accompanying nonresponse) strongly distorts the survey estimates (in this case, in the direction of politically conservative respondents). What can be learned for the future from this failure in the past is that, even with the massive amount of digitally supplied data described under the umbrella of 'big data', it is important to consider distortions that are almost certainly present before drawing conclusions based on only seemingly representative study populations.

However, in quantitative linguistic studies and other fields of empirical research, it is sometimes simply impossible to apply a probability sampling technique. Therefore, a non-probability sampling technique with the assumption of a realistic selection model can serve as a trade-off between fantasy and reality (Dörnyei 2007: 98). Yet the most important advantage of the probability over the non-probability sampling approach with respect to the representativeness of survey results is that the former can be applied, even when very little is known about the finite population, because the design unbiasedness of estimators holds in every case. In the strictly model-based approach of non-probability sampling techniques, the presumed model is always under discussion. To change the underlying model might yield different results. Therefore, the chosen model might not be generally accepted in cases where reliable statistics for people with very different opinions are needed (Thompson 2012: 131). Nevertheless, in practice, the design-based representativeness of probability sample results from survey research can also be seriously at risk for various reasons.

4. Errors that distort the representativeness of samples

The representativeness of samples (as defined in Section 3) refers explicitly to the so-called sampling error describing the sample-to-sample variation of sample estimators for population parameters under perfect conditions. However, the term also implicitly refers to so-called non-sampling errors in a survey. These are errors that can occur in both sample and population surveys. There are various possible sources of this category of errors in surveys. One of these sources is an incorrect sampling frame, which is the 'list' of individuals belonging to the population under study, for example, a certain language

community (Wolfram 2011: 297). The quality of the frame, as the basis from which a probability sample of population units can be drawn, depends on the relationship between the true population and its frame equivalent (Labov 1984: 47). For example, over-coverage will occur if units that are not part of the true population are included in the sampling frame. On the other hand, under-coverage will occur if units belonging to the true population are not included in the available frame population. The presence of over- or under-coverage in terms of possible representativeness of the sample can only be ignored if the coverage error is negligible. However, when frame imperfections are non-negligible, estimation techniques must be adapted regarding this target. For over-coverage, the theory of domain estimation may offer an appropriate approach. For under-coverage, the theory of ratio estimation may help to achieve (at least approximately) the goal of representativeness of survey results when the underlying model regarding the distribution of the variable under study in the non-covered subpart of the population of interest holds (Särndal, Swensson, and Wretman 1992: 544). Simply (re)defining the population under study as being the frame population may also be an option for producing representative samples of populations when frame imperfections occur. Clearly, the results then refer only to this frame population, and not, for example, to a population of residents.

In the presence of nonresponse, another source of non-sampling errors is the unavailability of survey units, their refusal to participate, or their noncooperation on certain items. The sample drawn from the population is therefore grouped into two sets: a response and a nonresponse set (Quatember 2015: 53–56). For sensitive subjects, such as poverty, sexual harassment, drug use, or membership in certain (generally unaccepted) subgroups of the population, the response set may be further divided into sets of truthful and untruthful answers. In this case, analyzing only the response set as if it were the original sample, an approach known as 'available cases analysis', will lead to a nonrepresentative sample of the characteristics under study if respondents differ from nonrespondents in this respect and/or if a 'socially desired' answering behavior occurs.

Clearly, the best way to circumvent these problems is to avoid both nonresponse and untruthful answering, because even the most sophisticated repair methods for the occurrence of nonresponse cannot be as good as observing true data. Scientists in the field of empirical social research have been considering the features of survey design that affect the quality and quantity of responses. Groves et al. (2004: 189–195) described several such methods found in the relevant literature, among which are the number and timing of attempts to access the sample person, data collection period, interviewer workload, interviewer observations, sponsorship, prenotification, incentives, burden, respondent rules, interviewer introduction behavior, interviewer-household matching, interviewer switches, data collection mode, and persuasion letters. Regarding mode of data collection, for example, the applied strategy (such as face-to-face, mail, telephone, or web interview) must not have a negative impact on response rates, hence tailoring the mode to the purpose of the survey is of great importance toward achieving representativeness of survey results.

Alternative indirect questioning designs have the potential to decrease rates of both nonresponse and untruthful answering, in particular for sensitive questions, by increasing respondents' privacy protections. Such questioning techniques can enable the representativeness of survey results, whereas a direct questioning technique might produce higher

nonresponse and untruthful answering rates (Chaudhuri and Christofides 2013; Quatember 2015: 85–120).

In the presence of nonresponse, two statistical approaches can be used to estimate the population characteristics of interest. First, one may estimate the characteristics based strictly on the responses by weighting adjustment. Second, the missing values themselves must first be estimated through a technique called data imputation. For both approaches to be efficient at achieving representativeness of survey results, meaning that the distribution or parameter of interest can be unbiasedly estimated (at least approximately), and that such an estimation satisfies a predefined accuracy requirement, it is necessary to formulate an adequate model of the mechanism that led to the occurrence of nonresponse. This means that the (experienced) user must have an idea to explain this process. For this purpose, Little and Rubin (2002: 11–12) distinguished between three possible types of nonresponse mechanisms: missing completely at random (MCAR), missing at random (MAR), and not missing at random (NMAR). The first term refers to a model in which nonresponse is assumed to occur completely at random. The second refers to a model in which response behavior depends only on observable information. The final refers to a model in which nonresponse depends on the missing information itself, and is therefore difficult to compensate for.

If, for instance, one wants to estimate the proportion of a certain language attribute in a simple random sample with nonresponse occurrence solely on the basis of the given responses, in the estimation process (Lohr 2010: 340–346), the original design weights have to be increased under an assumed nonresponse model because the observed respondents have to represent more units than desired. Under the MCAR model, for instance, the overall response rate can serve as an adequate adjustment factor to be applied to the design weights. Under the MAR assumption, the post-stratification procedure, in which the sizes of different subgroups of respondents are corrected to the known relative sizes of these groups in the population, can be useful. Again, representativeness can be achieved only if the applied model holds, in this case for the nonresponse mechanism, so that the original weights can be adjusted appropriately. Within the NMAR assumption, a model cannot completely compensate for nonresponse. However, in this case, the application of the aforementioned methods might at least also help to reduce the severe nonrepresentativeness of a survey regarding the characteristics under study.

The other statistical method to repair the effects of nonresponse on the representativeness of samples imputes a substitute for each missing value of the variable of interest. This results in a sample of the same size as the intended one, but with additional inaccuracy due to the imputation process. Different imputation methods, such as mean or regression imputation, in which the substitutes are derived from the available information in the observed data under an assumed nonresponse model, can be used to determine these substitute values (Little and Rubin 2002: 69–72). Under the assumption of an MCAR or MAR nonresponse mechanism, respectively, these imputation techniques are either applied overall or differentially within certain sample subgroups. For this method, the representativeness of a sample regarding a distribution or parameter of interest also depends on the adequacy of the nonresponse model on which selection of the applied imputation technique is based.

In practice, however, both methods are applied successively, starting with data imputation to compensate for item nonresponse, and continuing with weighting adjustment to compensate for unit nonresponse (Quatember 2015: 62–63).

5. Not all samples have to be representative

This discussion on representativeness of samples shall by no means ignore the fact that there are sample surveys whose purpose is not necessarily to reach conclusions based on calculated sample results about population-level parameters or distributions (Quatember 2019: 5–6), but which have a more 'informative' purpose (for example, a survey of groups of residents about language interactions, with the aim of having a 'first look' at certain aspects of this subject in order to formulate further research questions). Whether or not the 'representativeness' of the sample results (in the sense defined in Section 3) is actually necessary for the respective survey, or whether some degree of representativeness needs to be achieved, depends solely on the actual goals of the study. The users of the sample methods, however, must disclose the purpose of the survey, and their claim to quality derived therefrom.

6. References

Bellhouse, David R.
1988 A brief history of random sampling methods. In Paruchuri R. Krishnaiah & Calyampudi R. Rao (eds.). *Handbook of statistics*, vol. 6, 11−14. Amsterdam: Elsevier Science Publishers.

Bethlehem, Jelke
2009 The rise of survey sampling: Discussion paper (09015). The Hague/Heerlen: Statistics Netherlands.

Chaudhuri, Arijit & Tasos C. Christofides
2013 *Indirect questioning in sample surveys*. Heidelberg: Springer.

Deville, Jean-Claude & Yves Tillé
2004 Efficient balanced sampling: The cube method. *Biometrika* 91(4). 893−912.

Dörnyei, Zoltan
2007 *Research methods in applied linguistics*. Oxford: Oxford University Press.

Gabler, Siegfried
1996 Repräsentativität von Stichproben. In Hans Goebl, Peter H. Nelde, Zdenek Stary & Wolfgang Wölck (eds.), *Kontaktlinguistik: Ein internationales Handbuch zeitgenössischer Forschung*, 733−737. New York: De Gruyter.

Gabler, Siegfried & Andreas Quatember
2012 Das Problem mit der Repräsentativität von Stichprobenerhebungen. In Verband Schweizer Markt- und Sozialforschung (ed.), *Jahrbuch 2012*, 16−18. Zuerich: Verband Schweizer Markt- und Sozialforschung.

Groves, Robert M., Floyd J. Fowler, Jr., Mick P. Cooper, James M. Lepkowski, Eleanor Singer & Roger Tourangeau
2004 *Survey methodology*. Hoboken: Wiley.

Kruskal, William & Frederick Mosteller
1980 Representative sampling, IV: The history of the concept in statistics, 1895−1939. *International Statistical Review* 48(2). 169−195.

Labov, William
1984 Field methods of the project on linguistic change and variation. In John Baugh & Joel Scherzer (eds.), *Language in use: Readings in sociolinguistics*, 28−66. Englewood Cliffs: Prentice Hall.

Little, Roderick J. A. & Donald B. Rubin
2002 *Statistical analysis with missing data*. Hoboken: Wiley.

Lohr, Sharon L.
2010 *Sampling: Design and analysis*, 2nd edn. Boston: Brooks/Cole, Cengage Learning.
Milroy, Lesley & Matthew Gordon
2003 *Sociolinguistics: Methods and interpretation*. Oxford: Blackwell.
Neyman, Jerzy
1934 On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society* 97. 558–625.
Quatember, Andreas
1996 Das Problem mit dem Begriff Repräsentativität. *Allgemeines Statistisches Archiv* 80. 236–241.
Quatember, Andreas
2015 *Pseudo-populations: A basic concept in statistical surveys*. Cham: Springer.
Quatember, Andreas
2019 Inferences based on probability sampling or nonprobability sampling – Are they nothing but a question of models? *Survey Methods: Insights from the Field*. https://survey insights.org/?p=1011203. (accessed on 8 April 2019).
Särndal, Carl-Erik, Bengt Swensson & Jan Wretman
1992 *Model assisted survey sampling*. New York: Springer.
Squire, Peverill
1988 Why the 1936 Literary Digest Poll failed. *Public Opinion Quarterly* 52. 125–133.
Thompson, Steven K.
2012 *Sampling*, 3rd edn. Hoboken: John Wiley & Sons.
Weisberg, Herbert F.
2005 *The total survey error approach*. Chicago: The University of Chicago Press.
Wolfram, Walt
2011 Fieldwork methods in language variation. In Ruth Wodak, Barbara Johnstone & Paul Kerswill (eds.), *The SAGE handbook of sociolinguistics*, 296–311. Amsterdam: Elsevier Science Publishers.

Andreas Quatember, Linz (Austria)

43. The geolinguistic treatment of demolinguistic data

1. Introduction
2. Demolinguistics and geolinguistics
3. Gathering demolinguistic data for geolinguistic purposes
4. The geolinguistic treatment of demolinguistic data
5. References

1. Introduction

This chapter focuses on the geolinguistic treatment of demographic data, and more specifically of data elicited in multilingual settings. After establishing the basis of this chapter

https://doi.org/10.1515/9783110435351-043

by briefly situating and differentiating geolinguistics and demolinguistics as subdisciplines of linguistics (Section 2), we turn to a discussion of the kind of demolinguistic data on which a geolinguistic analysis is based and a brief presentation of the instruments (censuses, polls, and surveys) that are commonly used to gather such data (Section 3). The final section homes in on the geolinguistic treatment of demolinguistic data by means of language maps.

2. Demolinguistics and geolinguistics

A considerable amount of overlap exists between sociolinguistics, the sociology of language, contact linguistics, geolinguistics, and demolinguistics. Given this overlap, as well as the many competing broad and narrower definitions that are in use, it cannot be the purpose of this chapter to define these disciplines in detail, nor to arrive at a precise demarcation between the scope of demolinguistics as compared to geolinguistics, nor of these two as compared to sociolinguistics, the sociology of language, and/or contact linguistics (for a more detailed discussion, see Wildgen [2006] and Trudgill [2006]). For the purposes of this chapter it is, however, useful to at least broadly define the notions 'geolinguistics' and 'demolinguistics', since the labels are used in different ways in the literature.

On the one hand, geolinguistics refers to the study of the spatial distribution of specific (lexical, morphologic, phonetic, and/or syntactic) traits of one or more language varieties within a specific territory. In this sense, geolinguistics should be seen as a discipline that developed out of, and still has a lot in common with, dialectology (which emerged as a discipline at the end of the 19th century). The results of this type of geolinguistic research can be found in linguistic atlases all over the world (e.g. the French linguistic atlas, or ALF [Gilliéron and Edmont 1903–1910]). On the other hand, geolinguistics also refers to the spatial distribution of the use of language varieties, as well as the ruling beliefs and ideas (ideologies) about these varieties within a specific territory. In the words of Williams (1996: 63), it "seeks to illuminate the socio-spatial context of language use and language choice; to measure language distribution and variety; to identify the geographic characteristics of language groups in contact; to chart the dynamism of language growth and decline and to account for the social and environmental factors which create such dynamism" (see also Van der Merwe 1993: 411). This type of geolinguistics (or the 'geography of language', as some describe it) emerged along with other branches of hyphenated linguistics from the 1960s onwards. In this chapter, geolinguistics will be used as a synonym for the 'geography of language', and refers to the study of extralinguistic (rather than intralinguistic) features related to the distribution and use of, as well as attitudes about, one or more language varieties within a given territory (Dell'Aquila 2010). The focus on communitarian and social aspects of language varieties, as opposed to the specific linguistic performances of the users of such varieties, is a feature that geolinguistics (as used here) has in common with demolinguistics.

Some 25 years ago, De Vries (1990: 60–61) defined the term 'demolinguistics', and the kind of approach to the linguistic reality that it represents, in the following way:

> The term demolinguistics indicates that we are dealing with the application of one discipline (in this case, demography) to a specific object (in this case, language). Thus, we take an

> approach usually applied to total populations […] and adapted [to a linguistic community that] may be defined as a proper subset of an overall population, which is 'self-reproducing'. In other words, [it] must have a structure by age and sex of such a nature that (in principle, anyway) the group could reproduce itself by the normal demographic processes. […] A demolinguistic approach deals with the size, structure and geographic concentration or dispersion of language communities.

The classical demographic variables (e.g. age, sex, and birth or mortality rate) are meant to provide information on populations' stocks (i.e. the size of a particular population at a particular time) and flows (i.e. the changes to such populations as a consequence of entries and exits). Entries and exits refer to "the ways in which new members are added to the population and in which existing members are removed from it" (De Vries 1996: 771), following such processes as fertility, mortality, and migration. One can also decide to look at other modes of entry and exit that reveal information about certain changes in a certain subset of the population (a so-called quasi-population), such as, for example, a language community. For the variable 'language communities', there are not only the normal modes of entry and/or exit, but also an additional mode of entry, namely language shift (De Vries 1996: 772), which has an impact on the demographic weight of the quasi-populations that are in contact. Changes in the language behavior of a person belonging to a certain language community do not have an impact on the overall population in a certain setting. Migration from one language community to another, however, does have an impact on the distribution of the population of a language community. Information on language behavior is usually based on (census) data concerning the mother tongue, the vehicular language, or the official language in a certain area. As will be shown in the next section, geo- and demolinguistics are related, in terms of both data presentation and gathering.

3. Gathering demolinguistic data for geolinguistic purposes

Demolinguistic data for geolinguistic purposes are usually taken from at least one of the following sources: official population censuses, official language polls (taken at regular intervals), and large sociolinguistic surveys. Following De Vries (1990: 66), "[t]he most useful data source is a periodic count of the population, partitioned into linguistic quasi-populations by consistent criteria. Ideally, such counts should be obtained in a population census taken at regular intervals […]. Alternatively, one or more large surveys may provide a good starting point for analysis". It should be clear that the initial purpose for data collection is quite different in these cases: censuses and the like are obviously not devised to be primary sources of linguistic studies, and the variable 'language' which can be found there is but one of the variables − and not the main one − taken into consideration (Leeman 2017). The questions about languages quite often have a legal, political, or even ideological value, and are mainly employed for administrative purposes (see e.g. www.census.gc.ca for a list of topics covered in the Canadian Census).

This is different from the case of official language polls issued by governmental bodies, which specifically aim, for example, to gather data on the spread and use of certain languages and/or language varieties within (a part of) their specific territory. Sociolinguistic surveys that are not commissioned by governmental bodies tend to aim

at an even more detailed picture of the geographical distribution, as well as at the use of, and attitudes towards, language varieties in a certain area.

3.1. Census data

Questions about language started to be included in censuses from the mid-19th century in Belgium (1847), Switzerland (1850), British Ireland (1851), and Prussia (1852), followed by Austria-Hungary (1880), the USA (1890), British India (1891), and Russia (1897) (for details, see Duchêne and Humbert [2018], Duchêne, Humbert, and Coray [2018], and Leeman [2017]). Several countries have a long tradition of asking a limited number of language questions in censuses that not only allow them to monitor (changes in) the linguistic composition of the population, but also to adjust language policies and/or language-in-education policies on the basis of available data (Lieberson 1966). Questions about languages in censuses usually address (some of) the following topics: the population's 'mother tongue(s)', its 'language knowledge', and its actual 'language use' in certain domains. As explained by Arel (2004: 96), different censuses formulate questions related to these topics in quite different ways: "[f]or example, the 1890 Hungarian census defined mother tongue as 'that language which you recognize as your own and which you enjoy most speaking,' whereas the Prussian census defined it as 'that language in which one is most fluent from childhood on and in which one thinks and also prays'".

In what follows, language questions used in the Canadian, Finnish, and Swiss censuses will be briefly presented, since these censuses are generally considered accurate and reliable, both because of their inner structure and the care taken in data collection, and because of the high metalinguistic awareness of the population used to consider multilingual issues in their daily lives. In addition, some information is provided for the USA, South Tyrol (Italy), and India censuses, because they include some particularly interesting features.

The Swiss censuses of 1990 and 2000 included three language questions: (1) a question about the 'main language' (defined as the language one has best mastered and in which one thinks); (2) a question about 'home languages' (the languages usually spoken at home with relatives); and (3) a question about 'language use' at school and in the workplace. The respondents were given languages to choose from. The questions about 'home languages' and about 'school' and 'workplace' also included the option to select nonstandard varieties as an answer. As noted by Duchêne and Humbert (2018: 11), diglossia as well as linguistic differentiation are aspects "that are usually not quantitatively approached in official statistics". However, "[d]istinguishing languages from dialects or patois in censuses usually reveals a certain idea on the social cohabitation of varieties in opposition to a standard" (Duchêne and Humbert 2018: 11). It allows for real sociolinguistic and geolinguistic comparison based on a snapshot of the situation on the day of the census. The focus is on the actual inventory of the distribution of languages in the Swiss territory (not only the official languages, but also dialects and other [immigrant] languages), rather than on language policy assessment. Since 2012, the Swiss Federal Statistical Office conducts by means of a Computer Assisted Telephone Interview methodology a more in-depth thematic language survey based on a representative sample of roughly 16,500 people.

Of course, once available, census data can be used as background information to assess, adjust, or more broadly reflect on language policies that are in vigor (Kertzer and Arel 2004; Leeman 2017). Cases in point are Canada and Finland. The interest of the 2016 Canadian census (in its long form) lies in its high number (seven) of proposed language questions about: (1) knowledge of the official languages; (2) the language spoken most often at home; (3) languages spoken regularly at home; (4) respondents' mother tongue; (5) knowledge of nonofficial languages; (6) the language used most often at work; and (7) languages used regularly at work. These seven questions reveal the state's interest in the evolution of the linguistic repertoires of Canadian citizens, many of them having recently immigrated. The census results allow the government to reflect on educational provisions that, for example, simultaneously aim at maintenance of respondents' heritage languages as well as acquisition of the state languages (i.e. English and French).

From 1980, the USA census asks three questions: (1) whether the person speaks a language other than English at home; (2) what this language is; and (3) how well the person speaks English. These questions shift the main interest of the linguistic issue towards the diffusion of English in that society, with socio-political consequences (Duchêne and Humbert 2018; Leeman 2013).

On the other hand, in Finland (Finnish and Swedish are the official languages of the country) and in South Tyrol (officially bilingual Italian province), census responses are used for immediate administrative purposes. In Finland, they are, for example, used to define the official language of municipalities and the languages of instruction (on the basis of the results of the question about the mother tongue). In South Tyrol, they are used to regulate the 'linguistic composition' of administrative services according to languages typical of South Tyrol (on the basis of the results of the question about 'language group'). In both Finland and South Tyrol, the language data are enhanced by real-time information from registry offices, along with date and place of birth, sex, and so on. Finally, the census of India is also noteworthy; it provides quite detailed data about the vast linguistic diversity of the country, albeit with some difficulties due to non-standardized answers resulting from differences in popular and traditional names for the same language.

Further information on the censuses referred to above can be found on the websites listed below (last access on 12 September 2018):

- Canada: www12.statcan.gc.ca/census-recensement/index-eng.cfm
- Finland: www.stat.fi/tup/vl2010/tilastotietoa_en.html
- India: http://censusindia.gov.in/
- South Tyrol (Italy): http://astat.provinz.bz.it/de/default.asp
- Switzerland: www.bfs.admin.ch/bfs/en/home/statistics/population/surveys/volkszaehlung.html
- USA: https://www.census.gov/data.html

Once published, census data are of course at the disposal of both policy makers and linguistics researchers; despite being considered 'rough data', their use for validating pre-existing political or linguistic theories is nonetheless quite possible (Kertzer and Arel 2004; Leeman 2017).

3.2. Polls

While census data have the advantage of providing researchers with an amount of data that they would never be able to collect by themselves (i.e. data covering entire states or regions, or the possibility of a diachronic comparison), their disadvantage is that the language-related questions are not detailed enough to allow for more fine-grained sociolinguistic interpretations; indeed, they are actually not explicitly meant to serve for deep sociolinguistic analyses (Lieberson 1966). More detailed data can (and often are) gathered by means of large sociolinguistic polls that target a reduced population sample, and cover one or more limited geographical areas (usually well-established communities or multilingual regions). In the European realm, such polls are often organized, funded, and coordinated by regional governments or authorities in areas where so-called regional or minority languages are used by a part of the population, and where governments invest a quantity of (financial) means to maintain and/or revive the use of these languages. Examples include Catalonia, the Basque Autonomous Community, and Galicia in Spain; Friesland in The Netherlands; the Gaeltacht in Ireland (Ó Giollagáin et al. 2007), and South Tyrol in Italy. Outside of Europe, sociolinguistic polls have, for example, been publicly organized in Canada (the languages of the First Nations) and New Zealand (Maori).

In some cases, polls are administered entirely by departments in the administration of regional governments. In other cases, regional governments call upon the assistance of researchers based at research institutes or universities. Polls often contain several questions about language use, language competences, and language attitudes within a specific territory. The scope of the questions, and the difficulties related to asking questions in polls (and in nongovernmental sociolinguistic surveys, Section 3.3), will be broadly discussed below (Section 3.4). More detailed information on actual questions used in sociolinguistic polls is available on the following websites (last access on 12 September 2018):

- Basque Autonomous Community (Spain): www.euskara.euskadi.eus
- Catalonia (Spain): www.idescat.cat
- Friesland (The Netherlands): fryslan.gemeentedocumenten.nl/www.fryslan.frl/taalatlas.html
- Galicia (Spain): www.ige.eu
- Maori (New Zealand): www.stats.govt.nz
- South Tyrol (Italy): www.provincia.bz.it

3.3. Large-scale sociolinguistic surveys

Specifically centered on sociolinguistic issues are a few large-scale nongovernmental (even if they are still in many cases publicly funded) surveys of European regions. They include surveys in the Aosta Valley (Dell'Aquila and Iannàccaro 2003), Asturias (Llera Ramo 1994; Llera Ramo and San Martín Antuña 2003), the Ladin valleys (Dell'Aquila and Iannàccaro 2008; Iannàccaro and Dell'Aquila 2006), Latgalia (Dell'Aquila and Iannàccaro 2009; Iannàccaro and Dell'Aquila 2007), and Luxemburg (Fehlen and Piroth

1998). The aims of such quantitative surveys are either to gain coherent and complete knowledge about the current sociolinguistic profiles of the areas under analysis (e.g. by examining levels of use and knowledge rates of the spoken codes, as well as the social, ideological, and identity factors linked to the various languages), or to improve the sociolinguistic situation and vitality of a given code (normally perceived as needing special attention, which is, for example, the case for Ladin and Francoprovençal in Italy).

In any case, the large sample sizes and controlled sampling methods ensure a fairly high level of representativeness. It is clear that only careful, continuous, and methodologically rigorous comparison of census data and/or polls, and the outcomes of sociological and/or sociolinguistic research − which is more limited in space and time, but provides more detailed and precise language data − can provide a reasonably faithful picture of the use of a set of linguistic codes in a given area, and produce meaningful outcomes.

3.4. On questions posed and the kind of data obtained

As mentioned above, censuses contain a very limited number of questions about 'mother tongue' (or language first learned), '(official) language knowledge', and/or 'language use' (Arel 2004). Aiming at a more fine-grained picture of the linguistic state of affairs within a certain territory, polls commissioned by governments and sociolinguistic surveys obviously often contain a large set of more specific questions. The majority of questions is language-related, and meant to yield 'linguistic data', or data that can be used (socio)linguistically (see Iannàccaro [2000] on diverging opinions as to the nature of 'linguistic data').

Broadly speaking, the kind of data elicited can be categorized as cognitive data (number of language varieties known by respondents and their language skills), social data (use language varieties in different domains and their frequency), and affective or ideological data (attitudes, beliefs and ideas towards language varieties and on their social or territorial relationships). Of course "language census outputs cannot be dissociated from the process that makes them exist" (Duchêne, Humber, and Coray 2018: 46).

In order to be able to use these data to provide information about geographic and social variation within a given territory, questions that aim to gather nonlinguistic data (such as gender, birthplace, residence, occupation, and so on) are also part and parcel of surveys and polls. Many of them are similar to the kinds of demographic questions found in censuses.

A census is compulsory, and must be completed (either on paper or online) by the entire population. Polls and surveys are not compulsory. They rely on samples, and their level of representativeness is crucial to the reliability of data obtained. The elicitation methods can also be different. The methods that are most used are telephone interviews (although the data obtained are quite disputable) and written self-administered questionnaires; recently, online questionnaires have become more popular, but in this case, it is almost impossible to control for the validity of the sample.

An epistemological question that has often been debated is how realistic the data collected through self-administered questionnaires are − a bias with regard to censuses, polls, and surveys. As may be expected, ideological statements are more likely to emerge than objective descriptions. On the one hand, the researcher cannot verify the truth of the answers; on the other, the respondents may not be able (or even willing) to describe

their linguistic behavior in an objective manner. Such phenomena can be labelled psychological errors caused by some form of bias on the part of the respondent: typical outcomes of this type of error include overestimation of the diffusion of more prestigious languages (because respondents, either consciously or unconsciously, may feel obliged to claim knowledge of an official or a national language), and underestimation of the role of minority languages or varieties whose social status is low. Ideological biases should also be taken into consideration.

Respondent bias appears in different ways, depending on the types of question posed. A survey questionnaire normally opens with a few autobiographical questions (age, residence, religion, school, and so on), which are useful for establishing the demographic characteristics of the quasi-population polled; here, not counting the possibility of conscious lies from a respondent, the bias is normally quite low. More open to personal or ideological interpretations are the self-evaluating questions about language use in different domains (family, work, peer groups, and so on), and even more so, about language skills (listening, speaking, reading, and writing). Paradoxically, questions about metalinguistic judgements are less problematic, since there it is explicitly asked what the respondent thinks about the proposed situations (languages versus dialects, 'beautiful' or 'useful' languages, and so on); the same is true for questions about self-identification with territorial languages, or the status (or stigma) associated with particular codes. The less 'objective' the question, the more fitting, in this particular case, will the answer be. Other questions frequently asked center on judgements about potential linguistic policies which could be carried out. Given these constraints, particular care should be given to 'the art of asking questions' (Payne [1951] 2014); a detailed justification of the questions posed in a survey can be found in Iannàccaro and Dell'Aquila (2006).

Ultimately, what we really obtain is a declaration about 'opportunity of use', i.e. a good measure of the status and subjective position of the languages in the area under study, which allows us to ascertain the (socio)linguistic landscape of the speaker and her/his experience of the communities. Moreover, the wide range of questions in regional and local polls, together with a still highly representative sample, can actually provide a quite realistic overview of the current linguistic situation, both from a social and a geographical point of view.

When interpreting data that were not collected by a trustworthy and neutral entity, one also has to deal with intentional errors, that is to say, deliberate distortions of data by institutions or researchers for political or ideological reasons, either by simply falsifying the data (although this is a rare occurrence), or by designing survey questions or multiple-choice options in such a way as to influence informants' responses.

A meaningful geo-demolinguistic study should also interpret − and not just show − the data which are presented. This crucial statement may seem obvious, but it is often neglected. In this respect, the computer-aided treatment of large quantities of data (so-called data mining) can turn out to be an heuristic tool that not only leads to a better or more refined description of linguistic data, but also points to new ideas and suggestions for future research and even language policy actions (see for instance Dell'Aquila and Iannàccaro 2007; Iannàccaro 2011).

4. The geolinguistic treatment of demolinguistic data

Maps are one of the ways researchers transmit information that has previously been collected, selected, and analyzed (Goebl 1996; for more general information see Ambrose and Williams 1991; Dell'Aquila 2002; Dell'Aquila and Iannàccaro 2009; Iannàccaro and Dell'Aquila 2017; Kraak and Ormeling 2010); they depict graphically one or more phenomena or characteristics of a given reality by reference to a specific territory. In particular, a *geolinguistic* map for the representation of demolinguistic data has six main functions: to set the boundaries of the study; to observe, collect, and record information in an ordered manner; to store, retrieve, and update information consistently; to analyze by cartographic means; to present results; and to interpret these results (Ambrose and Williams 1991: 300).

The purpose of a map is therefore to represent an order, differences within an order, or exceptions to an order. Within geo-demolinguistic studies, 'representing an order' means mapping the intensities or percentages of use of a given linguistic code when these are subject to variation. Of course, variation constructed upon demolinguistic data may be diatopic, as it occurs along language borders, in linguistic enclaves, among ethno-linguistic minorities, and/or in any situation of linguistic contact. Variation may also be diachronic, although still diatopically defined, and in a certain sense, still comprehended as a diatopic representation (e.g. a map showing decreases in a minority language in favor of a dominant one); it may also mean graphically representing the vitality of a linguistic code, or its degree of officiality. Graphically portraying exceptions may mean highlighting small minorities or linguistic islands on the map.

A map is made to be looked at: it should attract attention and, at the same time, provide information. The added value of a map is that it bears the force of an image: in contrast to a text, which can only be read word-by-word, a map's fruition takes place simultaneously along two different dimensions. Yet while an argumentative text can offer sequential explanations of the phenomena studied − which are somewhat 'guided' by the text's author − a map provides access to a single piece (or several pieces) of information simultaneously, at the same time apparently leaving to its reader ample freedom in interpretation. Nonetheless, maps by their very nature must present information preorganized and filtered by the researcher, even if they seem to allow a non-guided interpretation. Each map, then, must be designed to present an autonomous aspect of reality, a kind of taxonomic representation (or a 'suggestion', as it were) consigned to the readers to analyze in their own personal way. Given its non-argumentative mode of representation, a map forces the reader to select information, focus on a single aspect, and rationalize their approach, in order to obtain and discuss research data (Dell'Aquila 2010).

Therefore, while in a textual, written discussion it is possible to compare and contrast different sets or sources of data in terms of their epistemological value and power to represent reality, when a map is at stake, the work of taxonomization and systematization must be done prior to its construction. The categorization of information into separate sets is implicit in all forms of scientific explanation, but is particularly marked in the interpretation of cartographic output. It is generally accepted that − without prejudice to the theoretical freedom pertaining to categorization of scientific information − the 'boxes' (types) in a grid used to interpret phenomena (tokens) must be inferior in number to the phenomena themselves by at least a couple orders of magnitude to enable grouping and comparison. However, thematic cartography entails an even narrower and more con-

centrated focus, necessarily excluding all opportunity for vertical (or paradigmatic) comparison among categories, and imposing a strictly syntagmatic perspective and analysis (Iannàccaro and Dell'Aquila 2017).

It is in relation to the peculiar circular dialectic between the two poles of theory and data that preparing a map becomes particularly challenging: maps − whether or not this aspect of them is recognized and discussed − are invariably a schematic, synthetic, and simplified interpretation of reality. It is critical to note here that extracting one's data from among the phenomena encountered necessarily involves making choices and applying exclusions: by definition, part of the information theoretically available to us does not fit in the 'boxes', and must therefore be overlooked, and even more so when it comes to mapping data. As a result, one of the most theoretically consequential means by which the taxonomization is accomplished consists of segmenting information (which is by nature contiguous) into discrete sets. Clearly, contiguous information is not only unmanageable, but essentially unknowable; therefore, in order to be presented and discussed as data, it must be assigned to discrete epistemological 'containers'.

Precisely because of this process of selecting and segmenting reality, a language map must be collocated in a wider graphical 'discourse'. Scientific explanation that is based on argument offers holistic and overarching perspectives on the phenomena under study by simultaneously presenting various categories and the relationships among them. If a single map, which analyzes one (or a limited number of) phenomena at a time, is to have explanatory power, it must be understood as representing a standalone aspect of reality, as though in a certain sense each individual map were a mini-classification theory. Therefore, in order to build up a complex picture of a given reality, a set of different maps and types of maps will be required. In practice, we need several different maps: they should not only be based on different types of data, but also be created by following a range of different methods.

Viewing different data in relation to the same basic map facilitates more thorough and deep-ranging inquiry. This allows the linguist to shed light on key aspects of linguistic situations, such as the concentration of speakers of a language in homogeneous areas, or spatial factors in language innovation and conservation, with the ultimate aim of building up the broadest possible picture of the linguistic identity of the area under study. Thus, "maps shape an imagined reality which takes on a life of its own and often become the basis for understanding and action" (Williams 1996: 66).

As such, language maps do not consist of a mere descriptive representation, but must be considered a fully heuristic tool for linguistic analysis. Cartographical representation of demolinguistic data therefore consists of the elaboration of sets of thematic maps, each of which represent a particular aspect of the (demo)linguistic reality in a way that makes them easily comparable.

5. References

Ambrose, John & Colin H. Williams

1991 Language made visible: Representation in geolinguistics. In Colin H. Williams (ed.), *Linguistic minorities, society and territory*, 298−314. Clevedon: Multilingual Matters.

Arel, Dominique
2004 Language categories in censuses: Backward- or forward-looking? In David I. Kertzer & Dominique Arel (eds.), *Census and identity: The politics of race, ethnicity, and language in national* censuses, 92–120. Cambridge: Cambridge University Press.
Breton, Roland
1976 *Atlas geographique des langues et des ethnies de l'Inde et du subcontinent: Bangladesh, Pakistan, Sri Lanka, Nepal, Bhoutan, Sikkim*. Quebec: Les Presses de l'Université Laval.
Bruk, Solomon I. & Vassilij S. Apenčenko
1964 *Атлас Народов Мира*. Moskva: Akademija Nauk SSSR.
Chambers, J. K. & Peter Trudgill
1980 *Dialectology.* Cambridge: Cambridge University Press.
Dell'Aquila, Vittorio
2002 Dati linguistici e cartografia tematica. In Roland Bauer & Hans Goebl (eds.), *Parallela IX*, 149–171. Wilhelmsfeld: Gottfried Egert Verlag.
Dell'Aquila, Vittorio
2010 GIS and sociolinguistics. In Alfred Lameli, Roland Kehrein & Stefan Rabanus (eds.), *Language and space: An international handbook of linguistic variation*, 458–482. Berlin & New York: de Gruyter.
Dell'Aquila, Vittorio & Gabriele Iannàccaro
1999 Elementi per lo studio delle frontiere linguistiche in val di Fassa. *Géolinguistique* 8. 5–49.
Dell'Aquila, Vittorio & Gabriele Iannàccaro
2003 Investigare la Valle d'Aosta: Metodologia di raccolta e analisi dei dati. In Rita Caprini (ed.), *Studi offerti a Michele Contini*, 221–243. Alessandria: Edizioni dell'Orso.
Dell'Aquila, Vittorio & Gabriele Iannàccaro
2007 Metodi statistici per la misurazione del plurilinguismo sociale e dei rapporti tra i codici. In Jeroen Darquennes (ed.), *Contact linguistics and language minorities / Kontaktlinguistik und Sprachminderheiten* (Plurilingua XXX), 77–89. Bonn: Asgard.
Dell'Aquila, Vittorio & Gabriele Iannàccaro
2008 Misurare il plurilinguismo: comunità e lingue nelle valli dolomitiche. In Gabriele Blaikner-Hohenwart, Evelyn Botolotti, Rita Franceschini, Emese Lörincz; Leander Moroder, Gerda Videsott, & Paul Videsott (eds.), *Ladinometria: Festschrift für Hans Goebl zum 65. Geburtstag / Miscellanea per Hans Goebl per il 65° compleanno / Publicazion en onour de Hans Goebl en gaujion de si 65 agn*, 229–258. Salzburg-Vich: Universität Salzburg, Freie Universität Bozen, Istitut Cultural Ladin "Majon di fascegn", Istitut Ladin "Micurà de Rü".
Dell'Aquila, Vittorio & Gabriele Iannàccaro
2009 Cartographical tools for sociolinguistic analysis: The survey Latgale. In Sanita Lazdiņa & Ilga Šuplinska (eds.), *Valodas Austrumlatviā: Pētītma dati un resultāti / Languages in Eastern Latvia: Data and results of survey*, 239–272 & 327–466. Rēzekne: Rēzeknes Augstskola.
De Vries, John
1986 *Towards a sociology of languages in Canada* (Publication B-153 of the CIRB). Laval: CIRB.
De Vries, John
1990 On coming to our census: A layman's guide to demolinguistics. *Journal of Multilingual and Multicultural Development* 11(1–2). 57–76.
De Vries, John
1996 Language demography. In Hans Goebl, Peter H. Nelde, Zděnek Starý & Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact linguistics / Linguistique de contact*, vol. I, 771–776. Berlin & New York: De Gruyter.

Duchêne, Alexandre & Philippe N. Humbert
2018 Surveying languages: The art of governing speakers with numbers. *International Journal of Sociology of Language* 205. 1–20.
Duchêne, Alexandre, Philippe N. Humbert & Renata Coray
2018 How to ask questions on language? Ideological struggles in the making of a state survey. *International Journal of Sociology of Language* 205. 45–72.
Fehlen, Fernand & Isabelle Piroth
1998 *Le sondage 'Baleine': Une étude sociologique sur les trajectoires migratoires, les langues et la vie associative au Luxembourg*. Luxembourg: SESOPI.
Gilliéron, Jules & Edmond Edmont
1903–1910 *Atlas linguistique de la France (ALF)*. Paris: Champion.
Goebl, Hans
1996 Sprachenkarten / Linguistic maps / Cartes linguistiques. In Hans Goebl, Peter H. Nelde, Zděnek Starý & Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact linguistics / Linguistique de contact*, vol. II, 1973–1980. Berlin & New York: De Gruyter.
Iannàccaro, Gabriele
2000 Per una semantica più puntuale del concetto di 'dato linguistico': Un tentativo di sistematizzazione epistemologica. *Quaderni di Semantica* 41(1). 21–49.
Iannàccaro, Gabriele
2011 Patterns of language maintenance: A quantitative approach. In Emanuele Miola & Paolo Ramat (eds.), *Language contact and language decay*, 137–163 & 249–283. Pavia: IUSS Press.
Iannàccaro, Gabriele & Vittorio Dell'Aquila
2006 *Survey Ladins: Usi linguistici nelle Valli Ladine.* Trento: Regione Autonoma Trentino-Alto Adige.
Iannàccaro, Gabriele & Vittorio Dell'Aquila
2007 Языки в восточной Латвии: методологические вопросы. In Ilga Šuplinska & Sanita Lazdiņa (eds.), *Acts of the international conference Etniskums Eiropā: sociālpolitiskie un kultūras procesi / Etniskums Eiropā: socialpolitiskī i kulturys procesi / Ethnicity in Europe: Sociopolitical and cultural processes*, 5–32. Rēzekne: Rēzeknes Augstskola.
Iannàccaro, Gabriele & Vittorio Dell'Aquila
2017 Sur la cartographie des données linguistiques, avec quelques considérations sur les langues en Crimée. In Marco Giolitto, Irina Ivanova, Sébastien Moret & Elena Simonato (eds.), *Les communautés suisses de Crimée et de la mer Noire : Langues et traditions* (Cahiers de l'ILSL 51), 91–138. Lausanne: Université de Lausanne.
Johnson, Leslie M. & Eugene S. Hunn (eds.)
2012 *Landscape ethnoecology: Concepts of biotic and physical space*. New York: Berghahn Books.
Kertzer, David I. & Dominique Arel (eds.)
2004 *Census and identity: The politics of race, ethnicity, and language in national censuses*. Cambridge: Cambridge University Press.
Kraak, Menno-Jan & Ferjan Ormeling
2010 *Cartography: Visualization of geospatial data*. Harlow: Pearson.
Leeman, Jennifer
2013 Categorizing Latinos in the history of the US Census: The official racialization of Spanish. In José Del Valle (ed.), *A political history of Spanish: The making of a language*, 305–323. Cambridge & New York: Cambridge University Press.
Leeman, Jennifer
2017 Censuses and large-scale surveys in language research. In Kendall A. King, Yi-Ju Lai & Stephen May (eds.), *Research methods in language and education*, 83–97. London & Berlin: Springer.

Lehmann, Christian
2004 Data in linguistics. *The Linguistic Review* 21. 175–210.
Lieberson, Stanley
1966 Language questions in censuses. *Sociological Inquiry* 36(2). 262–279.
Llera Ramo, Francisco J.
1994 *Los asturianos y la lengua asturiana: Estudio sociolinguistico para Asturias 1991*. Uviéu/Oviedo: Serviciu de publicaciones del Principáu d'Asturias.
Llera Ramo, Francisco J. & Pablo San Martín Antuña
2003 *II estudio sociolingüístico de Asturias*. Uviéu: Academia de la llingua asturiana.
Mackey, William F.
1988 Geolinguistics: Its scope and principle. In Colin H. Williams (ed.), *Language in geographical context*, 20–46. Clevedon: Multilingual Matters.
Meyerhoff, Miriam, Erik Schleef & Laurel MacKenzie
2015 *Doing sociolinguistics: A practical guide to data collection and analysis*. London & New York: Routledge.
Ó Giollagáin, Conchúr, Seosamh Mac Donnacha, Fiona Ní Chualáin, Aoife Ní Shéaghdha & Mary O'Brien
2007 *Comprehensive linguistic study of the use of Irish in the Gaeltacht: Principal findings and recommendations*. Dublin: The Stationery Office.
Payne, Stanley Le Baron
2014 [1951] *The art of asking questions: Studies in public opinion*, 3rd edn. Princeton: Princeton University Press.
Pike, Kenneth Lee
1967 *Language in relation to a unified theory of structure of human behavior*, 2nd edn. The Hague: Mouton.
Pop, Sever
1951 *La dialectologie: Aperçu historique et méthodes d'enquêtes linguistiques*. Louvain: Chez l'auteur.
Provinsje Fryslân
2007 *De Fryske taalatlas: Fryske taal yn byld*. Leeuwarden: Provinsje Fryslân.
Trudgill, Peter
1983 *On dialect: Social and geographical perspectives.* Oxford: Basil Blackwell.
Trudgill, Peter
2006 Sociolinguistics: An overview. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds.), *Sociolinguistics*, vol. 1, 1–5. Berlin & New York: De Gruyter.
Van der Merwe, Izak
1993 The urban geolinguistics of Cape Town. *Geojournal* 3. 409–417.
Wildgen, Wolfgang
2006 Sprachkontaktforschung. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds.), *Sociolinguistics*, vol. 2, 1332–1345. Berlin & New York: De Gruyter.
Williams, Colin H.
1988 An introduction to geolinguistics. In Colin H. Williams (ed.), *Language in geographical context*, 1–19. Clevedon: Multilingual Matters.
Williams, Colin H.
1996 Geography and contact linguistics. In Hans Goebl, Peter H. Nelde, Zděnek Starý & Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact linguistics / Linguistique de contact*, vol. I, 63–75. Berlin & New York: De Gruyter.

Vittorio Dell'Aquila, Milan (Italy)
Gabriele Iannàccaro, Milan (Italy)

44. Domain analysis

1. Introduction
2. Macro-dimension
3. Micro-dimension
4. Diachronic dimension
5. Normative dimension
6. Methods of domain analysis
7. References

1. Introduction

Domain analysis is an umbrella term for a range of approaches to the study of language use in different fields of human activity, generally termed *domains*. The notion of the domain is intuitively rather straightforward: it is a social setting, such as family, school, church, or business, in which one or more languages are used. The scope and meaning of a *domain analysis* depends to a great extent on the purposes for which it is conducted, and which dimension of language it involves. There are four broad dimensions of language: micro, macro, diachronic, and normative, which are closely interrelated (see Figure 44.1):

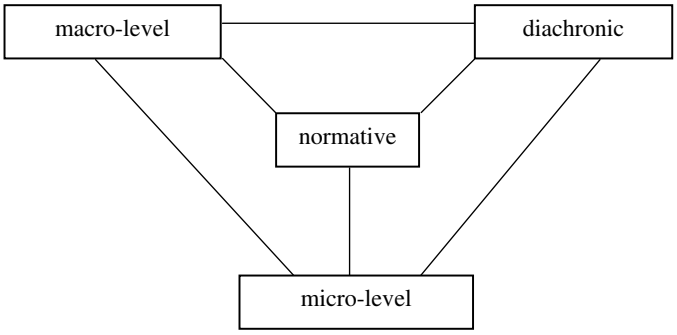


Fig. 44.1: Dimensions of language

On the macro-level, the goal of domain analysis is to specify the factors that regulate the use of one or another language in different domains. The choice of language by a speaker in any particular communicative act in any particular domain is called *code choice*. Code choice is a micro-level phenomenon. It depends partly on: domain-specific norms, the nature of the social networks with which the speaker is connected, conversational factors, and personal preferences. Micro-level code choices by individual users accumulate to shape the sociolinguistic norms of 'proper use' at the macro-level, which in turn guide the very same code choices at the micro-level.

In the diachronic dimension, a central notion of domain analysis is *domain loss*. Domain loss is a gradual replacement of one language with another in a given domain as a consequence of changes in code choices. Consecutive losses of domains may lead to language extinction. The real or imagined consequences of domain loss are the main

https://doi.org/10.1515/9783110435351-044

motivating factors for activities in the normative dimension of language. The normative dimension involves all meta-level practices − language policies, ideologies, and codifications − that purposefully influence the use of different languages in various domains. For example, language policy is a normative practice that seeks to shape the processes of code choice, either to secure the dominant position of one specific language in relation to its possible competitors (avoiding domain loss), or to promote the use of a language in new functional domains (language maintenance and revitalization).

In the following sections, the facets of domain analysis are outlined in more detail for each dimension. The chapter ends with an overview of the main methods for the study of code choice and the factors that affect it.

2. Macro-dimension

The first studies addressing institutionalized restrictions on language choice in different functional domains were conducted before WWII by the German scholar Schmidt-Rohr (1933), who analyzed the relationship between the use of German (dialects) and the national standard languages in countries where German was spoken as a minority language (Italy, Hungary, Belgium, and Switzerland).

A quarter of a century later, Ferguson (1959) provided an in-depth analysis of *diglossia*: a phenomenon in which two varieties of the same language are used in different domains. Diglossia is rooted in the differential status between two varieties: high (H) versus low (L). The H variety is the literary standard, while the L is the oral vernacular. The distinction between L and H is maintained by natural language acquisition in the family, where the L variety is learned, and in formal education, where the H variety is learned (Ferguson 1959). The fact that H is not the first language for any group of speakers distinguishes diglossia from other settings in which the standard language and dialects are in complementary use (Hudson 2002).

Diglossia emerges typically in cultures in which the classic literary texts embody the core values of the society, and so become sacralized. Therefore, the choice of H and L is strictly regulated. For example, in 1941, a Greek professor who misused the L variety of Greek in a scholarly publication was suspended for two months for harming the Greek language, and therewith the Greek nation (Frangoudaki 2002: 106). Because of the strong connection to core values of the society, diglossia tends to be relatively stable: Arabic diglossia has existed as long as there has been knowledge of Arabic; in other cases, diglossia has lasted at least a couple centuries (Ferguson 1959).

The similarity between diglossia and certain bilingual settings motivated Fishman (1967) to broaden the notion for any two languages, related or not, that were restricted to different domains. For Fishman (1967: 34), “bilingualism is essentially a characterization of individual linguistic behavior whereas diglossia is a characterization of linguistic organization at the socio-cultural level”. He outlined a typology of four classes: diglossia with bilingualism, diglossia without bilingualism, bilingualism without diglossia, and neither bilingualism nor diglossia.

Diglossia with bilingualism occurs when a single speech community uses two varieties of the same language, or two unrelated languages, in different functional domains. The classic cases of diglossia, such as in Arabic or Swiss German, belong to this class.

This type also involves cases in which two unrelated languages are in complementary use. For example, in Paraguay, Spanish and Guarani are spoken by the vast majority of society, so that educational, business, and formal matters are discussed in Spanish, and colloquial and intimate conversations are conducted in Guarani (Rubin 1985). Some other well-known cases are Yiddish and Hebrew, and Welsh and English. Fishman (1967) notes that any language in which there are different speech registers which are kept functionally distinct can be viewed as having both diglossia and bilingualism, if the notion of bilingualism is broadened to include competence in two varieties that have only subtle differences.

The second type, diglossia without bilingualism, occurs in settings involving two speech communities: the educated and wealthy elite which speaks one language, and the uneducated and poor majority which speaks another language. Since the two speech communities have very little interaction, and also operate in different domains, the segregation of codes in the corresponding domains is fairly strict. Diglossia without bilingualism was very common in Eastern Europe in the 18th and 19th centuries. For example, in Latvia and Estonia, the elite spoke Baltic German, while the peasant majorities spoke Latvian or Estonian. According to Fishman (1967), such settings are prone to demands for linguistic equality for the L variety, or secessionism. The struggle for independence of the Baltic nations is a good historical case of this, while the suppression of Catalan under Franco's regime, and as recently as 2010, is one of the factors that has fueled Catalonian secessionism (Connolly 2013).

The third type, bilingualism without diglossia, involves minority communities which have become fluent in the majority language, and have started to use it along with their heritage language in domains in which only the heritage language was previously used. Bilingualism without diglossia is a landmark of language shift, because the prestige of the majority language tends to result in a decrease in use of the minority language over time. According to Fishman (1967), the competition between the two varieties, which arises from the fact that there is no strict segregation between the two into different domains, renders this setting inherently unstable.

The fourth type, neither bilingualism nor diglossia, can occur, according to Fishman (1967), only in small isolated speech communities in which there is no role differentiation and everyone has access to all functional domains. However, as even the smallest communities have some rituals and settings open only to selected members, it is unlikely that a situation without any functional differentiation between speech registers ever actually occurs. Therefore, the setting of neither bilingualism nor diglossia tends to be self-annihilating, according to Fishman.

While Fishman mentioned only the ideally homogenous monolingual speech community, there is one more setting that could qualify as neither bilingualism nor diglossia: it is the setting in which speakers of different speech communities are rapidly brought into contact with each other. There is a need to communicate, yet there is no bilingualism, not to mention any rules for diglossia. This is a situation where pidgins and creoles arise, and the 'metrolinguistic setting' (Otsuji and Pennycook 2009) also belongs here. Both of these settings tend to be self-annihilating once the population stabilizes.

Fishman's expansion of the notion of diglossia to any two languages in complementary use has multiplied the number of settings that could be described as diglossic. As a consequence, the strict separation of H and L domains Ferguson described in his examples is often violated, in what is described as 'leaky diglossia' (Fasold 1984). The 'leak-

ing' between H and L is largely dependent on regimes of authority which value different ways of speaking. For example, even in the case of Arabic, there are different intermediate varieties between the H and L: the 'vernacular standards' that manifest their own linguistic domination (Eisele 2013).

Interactionalist sociolinguistic studies have shown how speakers actively negotiate code choices in conversation, instead of rigidly following the structural constraints of diglossia. Therefore, the early structuralist approach to diglossia has come to be seen as static and lacking an understanding of the political and ideological motivations behind the stratification of varieties (Jaspers 2017). However, according to Hudson (2002), leaking diglossia is characteristic of settings in which H and L are in competition as indexes of two group identities, while diglossia tends to be more categorical in settings in which the H variety lacks native speakers and an associated collective identity. In either case, the distribution of varieties in domains cannot be completely analyzed at the macro-level alone, but rather the actual code choices of language users in daily practice must be studied empirically at the micro-level.

3. Micro-dimension

The theoretical grounds for a methodologically elaborated domain analysis on the micro-level were laid out by Fishmann (1967, 1972). He hypothesized that code choice in bilingual communities is constrained by three factors − TOPIC, ROLE, and LOCALE − so that "only *one* of the theoretically co-available languages or varieties *will* be chosen by particular classes of *interlocutors* on particular kinds of *occasions* to discuss particular kinds of *topics*" (Fishman 1972: 437, italics in original).

Fishman (1972) did not specify beforehand the domains, but claimed that they should be discovered by analyzing the language choice(s) of individual communicators in different relationships communicating in different settings about different topics. To prove the construct validity of the notion of domain, Fishman (1972) reported data from several earlier studies on the use of Spanish and English in the Puerto Rican community in Jersey City.

For example, when subjects were presented with congruent clusters of TOPIC, ROLE, and LOCALE (such as the role of *parent*, the topic of *how to be a good son*, and the locale of *home*), and asked what language they would use in this domain, family and friendship domains emerged as significantly Spanish-speaking, and school and work domains as strongly English-speaking. No domain emerged as exclusively selecting only one of the codes, but overall, the differences in code choice in different domains were statistically significant. This non-categorical distribution of two languages indicates that the factors influencing code choice in any particular verbal interaction depend on more factors than just TOPIC, ROLE, and LOCALE.

Several approaches have focused on other sets of contextual factors involved in code choice. Some contextual constraints on code choice are derived from the structure of the conversation itself. For example, in a bilingual conversation, the code of the immediately preceding utterance in a dialogue is a strong predictor for the choice of the code of the next utterance, whether by the same speaker or by the conversational partner. Thus, a code-switch is more likely to occur in initiating turns as opposed to in response turns

(Auer 1998). The choice of a single code, of code-switching, or of code-mixing is also dynamically used in a conversation to construct and lay claim to specific social identities (Auer 2005).

Another set of factors that influence the choice of code derive from the structure of social networks, and in particular whether a given network is closed or open (Blom and Gumperz 1972). *Closed networks* are those in which people have a well-defined group of people to whom they are connected and who all know each other well. *Open networks* are those in which not all people in the network know each other. Blom and Gumperz (1972) conclude that closed networks favor dialect use and associated local values, while in open networks, the use of the standard language is widespread, being connected to broader national values. This factor is particularly relevant for heritage language maintenance in regard to the extent to which closed networks become open by inclusion of new members who do not speak the heritage language, and use of the heritage language starts to decline to accommodate these new members (Ó Riagáin 1997).

In some societies, code choice is prescribed (at least in some domains) by the force of (language) law. For example, after the collapse of the Soviet Union, the Baltic countries introduced language laws to prescribe the use of state languages in domains in which Russian was previously used. According to Gal (1989), in societies prescribing codes by law or by a strong social norm, code choice symbolically stresses the status of different sociopolitical positions. In the Baltic countries, for example, the new language laws expressed the status reversal of the Russian-speaking populations that had enjoyed the high status of the Russian majority in the Soviet Union, but had been newly ascribed the status of minority in the emerging Baltic republics (Ehala 2017).

While all of the factors outlined above indirectly influence code choice, what ultimately determines the choice of code might be "the speaker intentions and calculations to optimize rewards" (Myers-Scotton and Bolonyai 2001: 2). According to Rational Choice Theory (Scott 2000), individuals interact with each other so as to increase social approval and recognition. The choice of code is thus a rational decision which is made by taking into account both the constraints of the situation and the linguistic resources at the disposal of the individual.

The role of speakers' agency has motivated Saxena (2014: 95) to introduce the notion of 'lifestyle diglossia', which develops "from below in the way individuals' everyday socio-cultural practices and projection of identities shape their language practices which may lead to language attrition". He shows how the preference for different lifestyles on the part of Malays in Brunei favors the use of Malay or English in different domains, even though both of the languages may be considered to be H varieties in Brunei.

There is no doubt that all of the factors which affect code choice at the micro-level (whether the desire to optimize rewards, accommodate the communicative partner, conform to a social norm, or express one's lifestyle identity) feed back into the macro-level, and may, under certain circumstances, bring about diachronic changes that manifest as domain loss, or ultimately, language shift and extinction.

4. Diachronic dimension

According to Fishman (1965: 77), domain analysis repeated in the same population over time "may be used to represent the evolution of language maintenance and language

shift in a particular multilingual setting". As language shift is an intergenerational phenomenon, diachronic study thereof requires a significant timespan, which is the main reason why longitudinal domain analyses are the exception to the rule in the field.

However, in some domains, changes in language use have been rapid enough for longitudinal analysis to be feasible. These studies focus on *domain loss* (or *functional shift*): a gradual change in patterns of code choice in a specific domain, leading to the expulsion of one language in favor of another. For example, since the late 1990s, there has been a tendency toward increased use of English in science, higher education, and commerce, which has come at the expense of national languages. For example, Kuteeva and Airey (2014) report that almost all Swedish PhD theses in the natural sciences are now written in English, due to increasing use of English in Swedish higher education. In the humanities, the corresponding figure is approximately 40 %. University staff and students explain this shift as being due to a lack of proper Swedish-language science terminology. Furthermore, work environments are increasingly shaped by globalization and internationalization, with large international enterprises using English as the *de facto* corporate language. However, the situation is actually more complex than this: the multinational workforce in such enterprises, communicating both face-to-face locally and electronically across borders, often engages in quite flexible multilingual practices (Haberland 2005; van Mulken and Hendriks 2015).

Theoretically, domain loss has been seen as the main outcome of *linguistic imperialism*, a phenomenon by which an internationally prominent language (such as English or Russian) takes over domains in which other languages were previously used (Ferguson 2007). It is feared that domain loss in science and higher education is likely to snowball, ultimately affecting the language of instruction in secondary (and lower level) schools, as parents might prefer lower level education in the language that is also used at higher levels. It is also believed that loss of domains will harm terminology, ultimately further increasing the differences in status between international and national languages, which in turn would make gradual language shift more likely (Phillipson and Skutnabb-Kangas 2017). Quite a number of smaller nations, such as Denmark (Danish Ministry of Culture 2003), Sweden (Mål i mun 2002), and Estonia (EKAS 2004), have implemented extensive programs to maintain the status and position of their national languages. Such actions belong to the normative dimension of language.

5. Normative dimension

Language policy is the central research paradigm involved in the normative dimension of language. It addresses the pattern of code choices made by language users in different functional domains, and aims to regulate which language or variety should be used in which domain.

Spolsky (2007) specifies domains as institutionalized spheres of language use. He does not define domains by analyzing the patterns of code choices, as proposed by Fishman (1972), but specifies them beforehand, as Schmidt-Rohr (1933) does. According to Spolsky (2007), each domain has its own language policy, which in turn has three interrelated aspects: practices, beliefs, and management. Language *practice* is the day-to-day communication in a particular domain, and reflects the cumulative code choices

made by speakers: "the real language policy", as Spolsky (2007: 3) puts it. *Beliefs* are the speakers' understanding of the status of different varieties, and their usefulness in different domains. *Management* involves all actions undertaken by any group or authority to influence code choice in a particular domain, such as state documents aimed at developing national languages (as referenced above).

Besides language policy, the normative dimension also includes *language codification* (or corpus planning). Language codification involves conscious efforts directed at developing the writing system, standardized grammar, and dictionaries for a given language. According to Haugen ([1972] 2001), language codification is imposed by governments and other social institutions, and thus is characteristic of modernity, whereas most of the world's languages remain uncodified. However, codification has been recurrent in several historically 'high' cultures, and is thus not only unique to modern times. And by all means, language codification is the way in which a typical situation of diglossia (such as the Arabic one) comes into being in the first place.

In broader terms, language codification also involves any other conscious efforts to shape language, which can occur even in oral cultures (mostly as a result of 'identity politics'). Kloss (1967) refers to languages created out of language planning as *Ausbau* languages. For example, Thomason (2007) argues that 16th century Portuguese was consciously shaped in order to make it more distinct from Spanish, and that speakers of Uisai (a Papua New Guinea language) switched the feminine and masculine markers of gender agreement in order to increase differentiation from neighboring dialects. All consciously induced changes in language, regulation of language use, and normative beliefs about languages (i.e. all meta-level shaping of language) belong to the normative dimension of language.

Domain analysis as formulated in the 1960s by Fishman, and its application in language policy studies, was designed for the context of a highly standardized national language in contact with a minority language or an oral vernacular. However, sociolinguistic settings since then have changed considerably, partly because of an increase in linguistic diversity in Western societies due to immigration, and partly because of postmodernist ideologies which have promoted multilingual and multicultural values (and thus discouraged national unity and monolingualism). To a large extent, these developments have eroded strict norms regarding code choices in different domains. Instead, as some scholars claim, a new pattern of language use called 'metrolingualism' (in which there are no clear boundaries between languages used in interaction) has emerged (Otsuji and Pennycook 2009).

Ideologically, this line of thinking aims at the deconstruction of the very concept of language, and of the understanding of multilingualism as a plurality of discrete languages. For example, Møller (2008) argues that language users in modern multilingual settings might not distinguish between different languages, but instead use all linguistic resources at their disposal indiscriminately in order to achieve their communicative goals. In such a situation, "it is not adequate to categorise this conversation as bilingual or multilingual, or even as language mixing, because all these terms depend on the separability of linguistic categories" (Møller 2008: 218), but rather what occurs here is the rejection of the traditional concept of language altogether. Theoreticians of metrolingualism, taking inspiration from queer theory (which aims to deconstruct the gender dichotomy), therefore claim that urban linguistic practices work towards the "rejection

of ortholinguistic practices and a production of new possibilities" (Otsuji and Pennycook 2009: 247).

Certainly, metrolinguistic practices pose a problem for domain analysis, which depends on the presumptive existence of distinct language codes. Furthermore, metrolingualism as a language ideology argues against language codification: the very normative process that gives rise to H and L dichotomy and the whole range of related phenomena that constitute the objects of study for domain analysis. Therefore, for domain analysis to remain a viable research paradigm, metrolingualism and its ideology must be accounted for. The main question here is whether metrolinguistic processes are in any significant way different from the processes of pidginization and creolization (see Xu and Deterding 2017).

According to Winford (2013), settings with little bilingualism among interacting populations are common in all cases involving rapid relocation of people (whether due to trade, slavery, and/or colonization). A frequent result of these social processes is the emergence of pidgin languages. If a population using a pidgin stabilizes in an area, the pidgin starts to expand into new functional domains, and can evolve into a creole, together with the formation of its own speech community which also sees it as a signal of collective identity (Ehala 2018). A similar emergence of new varieties and identities out of multilingual practices has been observed in the case of several urban vernaculars, such as Tsotsitaal and Isicamtho in South Africa (Winford 2013), and the hybrid Turkish-German variety used by minority youth who do not necessarily have a Turkish background (Auer 2005).

The emergence of pidgins and their evolution into creole languages therefore provides ample evidence that the fluidity, context dependence, and low codification which are characteristic of metrolinguistic practices are a recurrent pattern in rapid language-contact situations, and that the processes of structuring and codification start as soon as a collective identity emerges among people brought together through migratory processes. All of this indicates that Fishman's (1967) fourth type (neither diglossia nor bilingualism), when it emerges because of the rapid relocation of speakers, is a transitional setting that evolves into one of the three other types.

It is another matter altogether as to whether highly codified varieties are in some sense artificial and invented, and therefore in discordance with the natural, fluid order of language. Integrational linguists (Harris 2002) are correct in pointing out that linguistic communication forms a seamless whole with nonverbal communication, and that meaning is often very context-dependent. Because of this, communication is possible, even if the communicators do not share a code.

However, communication without shared codes is efficient only in settings providing rich immediate contextual information and involving rather restricted roles and topics (such as in a market). It is not possible in domains in which innovative ideas about any imaginable topic need to be conveyed without any immediate context (such as in science). For these domains, highly codified, autonomous language is needed which is able to convey messages without the aid of nonverbal communication and/or any surrounding contextual information available at the moment of the communicative act.

The answer to the question of whether standard languages are oppressive and should be deconstructed is thus straightforward: it is a political trade-off. The lower the level of codification and the higher the fluidity, the less exclusive is the domain for language users with no linguistic background. However, such a deconstructed language is usable

only in highly context-dependent, local domains for very narrow purposes. If a community wishes to operate in any domains in which abstract ideas are communicated (education, science, religion, etc.), a highly codified variety is required. These domains, and the variety used therein, are by necessity exclusive, because they presuppose a strong investment in time spent learning to become competent in the topics under discussion, and the varieties needed for such discussions.

Although issues of linguistic justice inherent in the stratification of English, standard national languages, and various minority and dialectal varieties in a globalized world are hotly debated topics (see May 2014), the question of whether diglossia is something that should be condemned as inherently oppressive goes beyond the scope of this paper. However, the discussion above is a reminder that scholars working on domain analysis, whether they argue for language maintenance and codification or against all forms of ortholinguistic oppression, are themselves engaged and participating in the normative dimension of language.

6. Methods of domain analysis

Currently, a full-fledged domain analysis has four interrelated goals. The first is an empirical description of the actual code choice pattern in each particular domain; this goal can be best achieved using quantitative methods. The second goal of domain analysis is to reveal beliefs about the status(es) of competing languages, and the strength of the correlations between beliefs and code choice(s); both quantitative surveys and qualitative methods can be used. The third goal is to identify the institutional arrangements that seek to regulate the use of competing languages in each domain studied; this is mostly achieved through document and discourse analysis. The fourth, and ultimate, goal of domain analysis is to assess diachronic trends in code choice in domains; this is possible only by engaging in longitudinal studies.

6.1. Methods for studying code choice

In the 1970s, domain analysis was extensively used to account for the maintenance of the Irish language. The results of these studies emphasized the importance of social networks in code choice (Ó Riagáin 1997). In particular, it appeared that use of Irish was strongly influenced by the number of network members who had low Irish proficiency or were English speakers (Ó Riagáin 1997: 34). His studies showed that the more open the network, the more easily are new members who do not speak the heritage language able to enter the network. In particular, changes in household composition (reflecting inclusion of non-speakers of the heritage language) have been seen as important aspects of network change affecting code choice.

To study code choice in social networks, particularly in the family and amongst friends, self-reported questionnaires ask respondents to describe their social networks and what languages they use in interactions with each particular member of their social network. The data are analyzed for various statistical relationships between the variables

(Ó Riagáin 1997). Detailed network analysis of code choice in the family is particularly useful for detecting signs of the disruption of intergenerational language transmission.

Critics of network analysis have pointed out that social network analysis is harder to conduct for domains in which network structure is more open and the number of interacting people is larger. In these settings, it is most important to obtain reliable information about the proportion of overall use of one or the other variety. There are different ways to approach this. One widely used method is self-reported questionnaires, in which subjects report which languages they use in particular domains (see Landry, Allard, and Henry 1996). To obtain a detailed assessment of variable use, Likert scales are often used, in which subjects can choose the answer that best describes the use of one (L_1) or the other (L_2) variety: e.g. (1) only L_1, (2) mostly L_1, (3) more L_1 than L_2, (4) L_1 and L_2 equally, (5) more L_2 than L_1, (6) mostly L_2, and (7) only L_2 (Ehala and Zabrodskaja 2014).

Likert scales for code choice are particularly suitable for settings in which only two languages are in competition, but can also be used for settings involving more languages. For example, the scale could be formulated so that speakers' default language is rated relative to *other languages* (for example *more in L_1 than in other languages*). After indicating the level of use of the L_1 and 'other languages', the subject is then asked to name the other languages.

In some domains, code choice can also be studied empirically. For example, code choice in the domain of higher education is often studied by counting the number of publications and dissertations written in one or the other language. As for written code choice, diachronic data are readily available. To the extent that the large corpora of electronic written communication have become available for researchers, it is possible to analyze actual written code choice, as well as diachronic trends in other domains. It is also clear that the notion of domain can be narrowed as required by the study; for example, blogs are readily available corpora for such studies, and code-switching in blogs is extensively studied.

6.2. Methods for studying conditioning factors

Together with analysis of actual code choice, an elaborate domain analysis also involves a concomitant analysis of factors that may have an impact on code choice. In this way, the analysis of code choice can be suited to studies that analyze the impact of conditioning factors.

According to Ó Riagáin (1997), language competence is one of the strongest constraints on code choice. Different individuals meet the requirements of the linguistic market differently; because of their habitus, they have different configurations of language competencies. Language competency is mostly studied using self-reported assessments, the scales of which can vary considerably, from categorical yes/no choices, to Likert scales assessing overall competency, to detailed questionnaires focused on each linguistic subskill (e.g. reading, writing, listening, and speaking).

Language political interventions and language management initiatives (such as language legislation, formal language policy documents, and any other normative regulations that aim to influence the choice of different codes in any particular domain) also have an impact on code choice, at least in the long term. While language political studies

are mostly conducted independently, their connection with the study of code choice is a good practice in domain analysis; studies of language competition in higher education have quite persistently used such connections (see Hultgren, Gregersen, and Thøgersen 2014).

The final broad conditioning factor of code choice is formed by language attitudes and beliefs. The study of a wide range of language attitudes together with code choice has a long history, particularly under the ethnolinguistic vitality framework (Yagmur and Ehala 2011). This is mostly achieved by combining self-reported questionnaires of code choice with questionnaires of language attitudes. This method enables statistical analysis of relationships between code choice and attitudes. As quantitative analyses of attitudes and beliefs provide only very basic data about the phenomenon, a more thorough level of analysis is often included by adding a qualitative analysis of interview data in order to reveal deeper ideological frameworks that justify and explain particular language attitudes (McEntee-Atalianis 2011).

Overall, the analysis of code choice in domains, and the societal factors that influence it, is the main tool for studying language competition in an effort to understand stability and dynamics in multilingual settings, the processes of language maintenance and shift, and the emergence of new languages.

7. References

Auer, Peter
1998 *Code-switching in conversation: Language, interaction and identity*. London & New York: Routledge.

Auer, Peter
2005 A postscript: Code-switching and social identity. *Journal of Pragmatics* 37(3). 403–410.

Blom, Jan-Petter & John J. Gumperz
1972 Social meaning in linguistic structure: Code-switching in Norway. In John J. Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics: The ethnography of communication*, 407–434. New York: Holt, Rinehart & Winston.

Connolly, Christopher K.
2013 Independence in Europe: Secession, sovereignty, and the European Union. *Duke Journal of Comparative & International Law* 24. 51–105.

Danish Ministry of Culture
2003 *Sprog på spil: Et udspil til en dansk sprogpolitik* [Language at stake: A proposal for a Danish language policy]. Copenhagen: Danish Ministry of Culture.

Ehala, Martin
2017 After status reversal: The use of titular languages and Russian in the Baltic countries. In Michael S. Flier & Andrea Graziosi (eds.), *The battle for Ukrainian: A comparative perspective*, 563–585. Cambridge, Massachusetts: Harvard University Press.

Ehala, Martin
2018 *Signs of identity: The anatomy of belonging*. London: Routledge.

Ehala, Martin & Anastasia Zabrodskaja
2014 Hot and cold ethnicities in the Baltic states. *Journal of Multilingual and Multicultural Development* 35(1). 76–95.

Eisele, John C.
2013 Approaching diglossia: Authorities, values, and representations. In Aleya Rouchdy (ed.), *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, 21–41. London: Routledge.

EKAS
2004 *Development strategy of the Estonian language 2004–2010*. Tartu: Ministry of Education and Research.
Fasold, Ralph W.
1984 The sociolinguistics of society. Cambridge, MA: Basil Blackwell.
Ferguson, Charles
1959 Diglossia. *Word* 15(2). 325–340.
Ferguson, Gibson
2007 The global spread of English, scientific communication and ESP: Questions of equity, access, and domain loss. *Ibérica* 13(1). 7–38.
Fishman, Joshua A.
1965 Who speaks what language to whom and when? *La Linguistique* 2. 67–88.
Fishman, Joshua A.
1967 Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues* 23(2). 29–38.
Fishman, Joshua A.
1972 Domains and the relationship between micro-and macrosociolinguistics. In John. J. Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics: The ethnography of communication*, 435–453. New York: Holt, Rinehart and Winston.
Frangoudaki, Anna
2002 Comment: Greek societal bilingualism of more than a century. *International Journal of the Sociology of Language* 157. 101–107.
Gal, Susan
1989 Language and political economy. *Annual Review of Anthropology* 18. 345–367.
Haberland, Harmut
2005 Domains and domain loss. In Bent Preisler, Anne Fabricius, Hartmut Haberland, Susanne Kjaerbeck & Karen Risager (eds.), *The consequences of mobility*, 227–237. Roskilde: Roskilde University.
Harris, Roy
2002 *The language myth in western culture*. Richmond: Curzor Press.
Haugen, Einar
2001 [1972] The ecology of language. In Alwin Fill & Peter Mühlhäuser (eds.), *The ecolinguistics reader: Language, ecology, and environment*, 57–66. London: Continuum.
Hudson, Alan
2002 Outline of a theory of diglossia. *International Journal of the Sociology of Language* 157. 1–48.
Hultgren, Anna Kristina, Frans Gregersen & Jacob Thøgersen
2014 *English in Nordic universities: Ideologies and practices*. Amsterdam & Philadelphia: John Benjamins.
Jaspers, Jürgen
2017 Diglossia and beyond. In Ofelia Garcia, Nelson Flores & Massimiliano Spotti (eds.), *The Oxford handbook of language and society*, 179–196. New York: Oxford University Press.
Kloss, Heinz
1967 'Abstand languages' and 'Ausbau languages'. *Anthropological Linguistics* 9(7). 29–41.
Kuteeva, Maria & John Airey
2014 Disciplinary differences in the use of English in higher education: Reflections on recent language policy developments. *Higher Education* 67(5). 533–549.
Landry, Rodrigue, Real Allard & Jacques Henry
1996 French in South Louisiana: Towards language loss. *Journal of Multilingual and Multicultural Development* 17(6). 442–468.

Mål i mun
2002 *Mål i mun: Förslag till handlingsprogram för svenska språket* [Proposal for an action plan for the Swedish language] (SOU 2002:27). http://www.regeringen.se/rattsdokument/statens-offentliga-utredningar/2002/04/sou-200227/. (31 October 2017).
May, Stephen
2014 Contesting public monolingualism and diglossia: Rethinking political theory and language policy for a multilingual world. *Language Policy* 13(4). 371–393.
McEntee-Atalianis, Lisa
2011 The value of adopting multiple approaches and methodologies in the investigation of ethnolinguistic vitality. *Journal of Multilingual and Multicultural Development* 32(2). 151–167.
Møller, Janus Spindler
2008 Polylingual performance among Turkish-Danes in late-modern Copenhagen. *International Journal of Multilingualism* 5(3). 217–236.
Myers-Scotton, Carol & Agnes Bolonyai
2001 Calculating speakers: Codeswitching in a rational choice model. *Language in Society* 30(1). 1–28.
Ó Riagáin, Pádraig
1997 *Language policy and social reproduction: Ireland 1893–1993*. Oxford: Clarendon Press.
Otsuji, Emi & Alastair Pennycook
2009 Metrolingualism: Fixity, fluidity and language in flux. *International Journal of Multilingualism* 7(3). 240–254.
Phillipson, Robert & Tove Skutnabb-Kangas
2017 English, language dominance, and ecolinguistic diversity maintenance. In Markku Filppula, Juhani Klemola, & Devyani Sharma (eds.), *The Oxford handbook of World Englishes*, 312–332. Oxford: Oxford University Press.
Rubin, Joan
1985 The special relation of Guarani and Spanish in Paraguay. In Nessa Wolfson & Joan Manes (eds.), *Language of inequality*, 111–120. The Hague: Mouton.
Saxena, Mukul
2014 "Critical diglossia" and "lifestyle diglossia": Development and the interaction between multilingualism, cultural diversity, and English. *International Journal of the Sociology of Language* 225. 91–112.
Schmidt-Rohr, Georg
1933 *Mutter Sprache: Vom Amt der Sprache bei der Volkwerdung*. Jena: Eugen Diederichs.
Scott, John
2000 Rational choice theory. In Gary Browning, Abigail Halcli & Frank Webster (eds.), *Understanding contemporary society: Theories of the present*, 126–138. London: Sage.
Spolsky, Bernard
2007 Towards a theory of language policy. *Working Papers in Educational Linguistics (WPEL)* 22(1).
Thomason, Sarah G.
2007 Language contact and deliberate change. *Journal of Language Contact – THEMA* 1. 41–62.
van Mulken, Margot & Berna Hendriks
2015 Your language or mine? Or English as a lingua franca? Comparing effectiveness in English as a lingua franca and L1-L2 interactions: Implications for corporate language policies. *Journal of Multilingual and Multicultural Development* 36(4). 404–422.
Winford, Donald
2013 Social factors in contact languages. In Peter Bakker & Yaron Matras (eds.), *Contact languages: A comprehensive guide*, 363–416. Boston & Berlin: de Gruyter.

Xu, Zhichang & David Deterding
2017 The playfulness of 'new' Chinglish. *Asian Englishes* 19(2). 116–127.
Yagmur, Kutlay & Martin Ehala
2011 Tradition and innovation in the Ethnolinguistic Vitality Theory. *Journal of Multilingual and Multicultural Development* 32(2). 101–110.

Martin Ehala, Tartu (Estonia)

45. Language typology in contemporary perspective

1. Introduction
2. A geographical foundation
3. Disciplinary themes
4. A contemporary typological model
5. Applications and evaluations
6. A concluding remark
7. References

1. Introduction

Dominant languages increasingly press upon the traditional boundaries and domains of less powerful ones. While the most obvious and most widespread contemporary pressures are those exerted by English, there are other varieties whose circumstances – existing or unfolding – may create difficulties for neighbors near and far. Indeed, it is not at all uncommon to find languages that are simultaneously bearing down upon others while under pressure themselves. East African Swahili (the name for which derives from an Arabic term for 'coast', a clear indication of its original importance), is an official variety in Uganda, Tanzania, and Kenya, and acts as a *lingua franca* among speakers of many more regional varieties: Bemba, Maasai, Luganda, Luo, Kikuyu, and so on. Some of these are, themselves, link languages among populations speaking still 'smaller' ones. At the same time, Swahili is under considerable pressure from English. Other relevant examples can be found in the former Soviet republics, where (among analogies to the English-Swahili-Bemba continuum, for example) we find chains linking Russian-Georgian-Ossetian and Russian-Moldovan-Gaugaz. One might also include the English-French-Cree sequence in Québec, as well as the South American ones that bring together Spanish-Quechua-Aymara and Spanish-(or Portuguese-)Guaraní-Chiquitano.

I begin with these examples to reinforce the usefulness of classifying social and linguistic points of contact in a world in which simplistic dualisms of the 'larger-language-against-smaller-language' type are rarely accurate (Mufwene [2005] provides one of the few truly balanced discussions of the operation of linguistic predation). Of course, the cases I have just highlighted are themselves more linguistically complex than is shown here. The implication for present purposes is that comprehensive descriptions of languages in contact – of relevance to our understanding of language dynamics generally, and more specifically, to all aspects of language policy and planning – will benefit from some formalized framework that will guide research exercises and facilitate cross-

https://doi.org/10.1515/9783110435351-045

context comparisons. We come, then, to typological exercises. Reflecting a taxonomic intent, the term *typology* is used here to maintain verbal continuity with most of the earliest contemporary efforts in the area (some discussed below). It does not connote the classificatory attention given to language structures and functions that figure in variationist dialect studies, nor that which is closely concerned with the sociolinguistics of language change.

For obvious reasons, the most interesting (and generally most pressing) of such exercises bear upon interrelationships among 'smaller' and 'larger' languages. Attention to the consequences of these contacts has become increasingly frequent since Krauss (1992) described the perilous state of many varieties, and there is now a burgeoning literature on endangered and 'at-risk' languages. Growth here has been accelerated by globalization processes; these mean that important interactions among languages are no longer restricted to varieties in physical contiguity (if, in fact, this was ever strictly the case; for important discussions in this regard, see ELDIA [2010–2013, 2013], Lewis [2005], and UNESCO [2003, 2010]). Too few treatments, however, have gone beyond that inaccurate binarism noted above, or risen above a simple condemnation of linguistic 'imperialism' coupled with an equally blunt endorsement of its 'victims'. I use the two terms advisedly here to draw attention to the characteristic lack of nuance and precision found in jejune large-versus-small arguments (one is here also reminded of Russell's [1950] warning against seeing a necessary link between 'virtue' and some form or another of subaltern status). Given that some of the work on languages in contact is quite strongly polemical, it is good to acknowledge more balanced treatments: works by Grenoble and Whaley (1998a, 2006), Austin and Sallabank (2014), and Jones (2015) come immediately to mind.

The close connections between much typological work and matters of linguistic endangerment obviously imply attention to linguistic vitality (to which I turn below), a topic which has had an important place in the sociolinguistic literature for four decades. Brief contemporary overviews are provided by Ehala (2015) and Smith, Ehala, and Giles (2017); Moretti, Pandolfi, and Casoni (2011b) provide a good sense of current applications. While devoted mainly (but not exclusively) to Italian in Switzerland, the latter book contains several contributions of generalizable value, some of which have typological themes. Moretti, Pandolfi, and Casoni (2011a) thus discuss indicators of vitality of Swiss Italian (see also Moretti and Pandolfi 2011), Iannàccaro and Dell'Aquila (2011) deal with Ladin in the Dolomites, and Stähli (2011) treats Walser German in Ticino. Ammon (2011) presents a checklist of vitality indicators for diaspora minority groups; his work is reinforced and amplified − here and elsewhere − by that of Gawrisch (2011). Berruto's (2011) closing commentary summarizes the impact and implications of this useful collection.

Recent attempts at typological profiling in the sociology of language are formally justified by remarks made by (among others) Haugen and Ferguson. As Haugen (1972: 325) states, "most language descriptions are prefaced by a brief and perfunctory statement concerning the number and location of its speakers and something of their history. Rarely does such a description really tell the reader what he ought to know about the social status and function of the language in question". Ferguson (1991: 230) builds on this, stating that "[i]t is frustrating to read a stimulating case study and find that it lacks information on what the reader regards as some crucial points […] what I have in mind is not so much a well developed theoretical frame of reference as something as simple as

a checklist of points to be covered". In fact, a necessary underpinning of any meaningful description is a careful assemblage of relevant facts − 'relevant', because the inclusion of polemical material has often hampered clear understanding of linguistic situations, and often served only to support outbreaks of special pleading (recent examples can be found in much of the literature on the ecology of language; see Edwards 2008, 2009, 2010). Inquiries that make some attempt at disinterested assessment must aim for that scientific validity which balances comprehensiveness with the wielding of Occam's razor.

Such attempts would of course be of little value if each real-life instance of language contact were unique. As in other disciplines, social-scientific descriptions assume that relatively stable generalities can be established. It is true that each context differs to some degree from any other, but anyone who has attempted a contrastive analysis, or who has tried to abstract general points from a variety of linguistic scenarios, has in effect argued that some features are constant, or at least similar enough across contexts, to suggest useful generalization(s). My contention has always been that, like a kaleidoscope, the uniqueness of each setting derives from the particular constellation of features that are, in and of themselves, remarkably common.

2. A geographical foundation

An obvious ground on which to base any sociology-of-language taxonomy is that of physical geography. Anderson's (1981) categorization scheme was perhaps the most fully fleshed of earlier attempts here, and his seven-part framework − which focuses particularly on degrees of contiguity, and what might be styled the 'minorities-here-but-majorities-elsewhere' situation − remains of interest. He was, for instance, one of the first to discuss, however incompletely, the phenomenon of minorities within minorities. This has lately attracted more attention (see Eisenberg and Spinner-Halev 2005), and rightly so, for it throws into relief the inconsistencies that arise when one group (say, the Georgians or the sovereigntist Québécois) makes a case for increased autonomy while simultaneously denying the validity of the same argument made by internal populations (for example, the Ossetians or the James Bay Cree, respectively); see also above. Frequent historical experience teaches us that possession of some form of minority status does not necessarily sensitize such groups to the perceived plight of other entities. This is yet another argument, of course, for the careful construction of taxonomies and typologies.

In attempting to describe the Italian context, White (1987) provided a framework that I then used as the foundation of my own. His model rested upon three basic dimensions: (i) the minority-language group in question may (or may not) be a majority one elsewhere; (ii) within a region, speakers may be concentrated or scattered; and (iii) speakers of the same language may (or may not) live in close proximity across regional borders. Expanding on these three dimensions, and taking into account both indigenous and immigrant groups, led to the geographical framework shown in Table 45.1. Details of other geographical models, and of my expansion of White's, may be found in Edwards (2009), where limitations of and lack of precision in the framework are pointed out as constructive criticism of some of these frameworks. One should also be aware that the terms

Tab. 45.1: Geographical contexts (only minority-language situations are illustrated)

TYPE	*INDIGENOUS MINORITIES*	*IMMIGRANT MINORITIES*
1. Unique Cohesive	Sardinian (Sardinia); Welsh (Wales); Friulian (Friuli-Venezia-Giulia)	Dialect communities (often religiously organized) in which the variety is now divergent from that in the region of origin (e.g. Pennsylvania 'Dutch')
2. Unique Non-Cohesive	Cornish (Cornwall)	As above, but where speakers are scattered
3. Non-Unique Adjoining Cohesive	Occitan (Piedmont, Liguria, and France); Basque (France and Spain); Catalan (Spain and Andorra)	Enclaves of immigrants found in neighboring states
4. Non-Unique Adjoining Non-Cohesive	Saami (Finland, Norway, Sweden, and Russia)	Scattered immigrants in neighboring states
5. Non-Unique Non-Adjoining Cohesive	Catalan (Spain and Sardinia)	Welsh (Patagonia); Gaelic (Nova Scotia)
6. Non-Unique Non-Adjoining Non-Cohesive	Romany (throughout Europe)	Scattered immigrants of European origin in 'new world' receiving countries
7. Local-Only Adjoining Cohesive	French (Valle d'Aosta and France)	French (in New England town enclaves); Spanish (southwest USA); Italian *Gastarbeiter* (Switzerland)
8. Local-Only Adjoining Non-Cohesive	German (Piedmont and Switzerland)	French (scattered throughout New England)
9. Local-Only Non-Adjoining Cohesive	French (Apulia and France)	Immigrant enclaves in 'new world' countries
10. Local-Only Non-Adjoining Non-Cohesive	Albanian (throughout the Mezzogiorno and in Albania)	As above, but where speakers are scattered

'indigenous' and 'immigrant' are not without definitional difficulties – resting, most obviously (but not exclusively) upon temporal factors.

3. Disciplinary themes

Geographical beginnings must be supplemented with information from the several disciplines that bear upon real-world language contact, and a number of scholars have made important contributions. Ferguson (1962) pointed to the importance of degrees of domi-

nance and scope (written, standardized, and so on). Stewart (1962, 1968) emphasized language types (pidgins, vernaculars, classical, artificial, etc.) and functions (among them, restricted, official, wider communication). Kloss (1967) discussed the extent and prevalence of multilingualism, the legal status of languages, and linguistic prestige. Most notable, perhaps − at least in terms of continuing relevance and citation − was Haugen's (1972) treatment of language ecology. This naturally stressed the study of interactions between a language and its environment, and Haugen suggested a number of specific questions to be dealt with. These touched upon historical, demographic, philological, attitudinal, and other aspects. Haugen's insights were elaborated on by Haarmann (1986, 1996), who was particularly interested in the role played by various factors in matters of language maintenance and shift. The work of Giles and his associates (Giles, Bourhis, and Taylor 1977; Bourhis, Giles, and Rosenthal 1981) is also important here, particularly with their concept of *ethnolinguistic vitality*, which grew from a consideration of more or less objective factors (demography, institutional and official support, and status) to include the more 'subjective' (but no less important) vitality deriving from and reinforcing psychosocial perceptions. See again the fuller details in Edwards (2009), which include my comments and, especially, criticisms − all of which of course fed into the development of my own model. Other scholars have also weighed in with typological contributions, but none with the scope of those mentioned here. See Clyne (2003) for useful criticism of some of these frameworks.

4. A contemporary typological model

An initial consideration of variables that are important for the assessment of language status (variables which have been regularly and repeatedly stressed in the literature), led to the suggestion of three very rough and basic categories: *speaker*, *language*, and *setting*. These are not, of course, watertight and mutually exclusive compartments, but they may serve as logically important benchmarks. For example, it is possible to list all relevant variables under one or more of these three headings, and they do reflect the spirit of an ecological inquiry − that is, one that emphasizes the interactions between language and environment.

We could also consider, by way of cross-perspective, a categorization of different *disciplinary perspectives*. The following immediately suggest themselves as germane: demography, geography, economics, sociology, linguistics, psychology, history, politics-law-government, education, religion, and the media. Again, these are hardly mutually exclusive categories, nor do I suppose that these eleven cover all necessary ground.

A simple cross-tabulation of speaker, language, and setting variables with the disciplinary perspectives just noted gives rise to the sort of framework depicted in Table 45.2. While it is not difficult to think of the sorts of questions suggested by each of the 33 points of intersection (or 'cells'), the ones shown here are really only broad suggestions, and are not anywhere near specific enough, by themselves, to comprise a complete or usefully applicable typology: they are best thought of as points of departure. It is also clear, incidentally, that some questions could plausibly fit in more than one cell. With a reminder to readers of the preliminary nature of the undertaking, then, a list of sample questions (one for each cell), keyed by reference to Table 45.2, appears as Table 45.3.

Tab. 45.2: A sociology-of-language framework for minority (and other) languages, showing 33 points of intersection (or 'cells')

		CONSECUTIVELY NUMBERED 'CELLS'		
	DISCIPLINARY PERSPECTIVES	*SPEAKER*	*LANGUAGE*	*SETTING*
1.	Demography	1	2	3
2.	Geography	4	5	6
3.	Economics	7	8	9
4.	Sociology	10	11	12
5.	Linguistics	13	14	15
6.	Psychology	16	17	18
7.	History	19	20	21
8.	Politics-Law-Government	22	23	24
9.	Education	25	26	27
10.	Religion	28	29	30
11.	Media	31	32	33

Tab. 45.3: Some sample questions, keyed to the 'cells' shown in Table 45.2

Demography
1. Numbers and concentrations of speakers?
2. Extent of the language (see also geography)?
3. Ruralurban nature of setting?

Geography
4. Immigrant or indigenous minority?
5. Extent of the language (see also Demography)?
6. Language uniqueness, cohesion, and contiguity?

Economics
7. Economic health of speaker group?
8. Association between language(s) and economic success/mobility
9. Economic health of the region?

Sociology
10. Socioeconomic status of speakers?
11. Degree and type of language transmission?
12. Nature of previous/current maintenance or revival efforts?

Linguistics
13. Linguistic capabilities of speakers?
14. Degree of language standardization?
15. Nature of in and outmigration?

Psychology
16. Language attitudes of speakers?
17. Aspects of the language-identity relationship?
18. Attitudes of majority group towards minority?

History
19. History of the group?
20. History of the language?
21. History of the area in which the group now lives?

Politics-Law-Government
22. Rights and recognition of speakers?
23. Degree and extent of official recognition of language?
24. Degree of autonomy or 'special status' of the area?

Education
25. Speakers' attitudes towards and involvement in education?
26. School support for the language?
27. State of education in the area?

Religion
28. Religion of speakers?
29. Association between the language and religion?
30. Importance of religion in the area?

Media
31. Group representation in media?
32. Language representation in media?
33. General public awareness of area?

5. Applications and evaluations

The typological approach that I have briefly outlined here was initially sketched in earlier publications (Edwards 1991, 1992), and others have thus had a chance to consider, apply, and criticize it. A recurring point made by many scholars is that the typology requires modification − sometimes at a basic level, sometimes for quite particular circumstances; see Valiquette (1998), for instance. It is both impossible and inappropriate here to go into specifics, but I very much endorse further refinements of what − as I point out once again − remains an introductory effort. Grenoble and Whaley (1998b) draw centrally upon the model, for instance, and cite both strengths and weaknesses of it. They note that it usefully distinguishes between the speech community in question and the surrounding context, while at the same time emphasizing the intertwining of variables at all levels of specificity. On the other hand, they (rightly) reveal the need for further model elaboration: some terms are incompletely or inadequately defined, variables might profitably be placed in some hierarchical order, and more focused attention on literacy might well be highlighted. Clyne (2003: 244) also pointed to the need for greater descriptive clarity, while referring favorably to the contextualization of variables, something that "could be considered more in the methodology of future studies". Fuller contextualization is indeed a central thrust of the framework.

In her study of language shift and revival among Quechua speakers in Ecuador, King (2001) briefly discusses the model, citing it along with Fishman's (1991) Graded Intergenerational Dislocation Scale and a three-part framework suggested by Hyltenstam and Stroud (1996). The latter emphasizes the social conditions that surround languages (see the 'extralinguistic framing' model they present) and dictate their fortunes. The authors focus on variables having to do with the language community at both the social and individual levels, and with the sociology of majority-minority interaction(s). The approach is thus broadly similar in intent to my own, but it does not highlight matters at quite the same levels of specificity. Fishman's approach is much less useful, in that his eight-point scale of obstacles to revival represents only a formalization of the familiar challenges faced by 'small' languages. It is − to use his own comparison − a sort of Richter scale of endangerment. Furthermore, since Fishman's intent is to outline the stages by which minority-language shift can be reversed, the model is more of a hortatory action plan than a purely descriptive framework. As Clyne (2003: 64) observes, the steps towards the reversal of shift seem to not coincide very well with the desired life trajectories of many immigrant minority populations: "many of the measures suggested by Fishman would tend to detract from [their] socioeconomic mobility and would therefore not appeal to most".

In their studies of Bashkir, Altai, and Kazakh speakers in the Russian republics of Bashkortostan and Altai, Yağmur and Kroon (2003, 2006) have employed the framework in conjunction with the ethnolinguistic-vitality approach of Giles and his colleagues (Giles, Bourhis, and Taylor 1977; Bourhis, Giles, and Rosenthal 1981). Paulston et al. (2007) have referred to it in their examination of 'extrinsic' linguistic minorities − that is, groups who once belonged to a majority population in a neighboring country: Russians in Latvia, being the clearest case in point. "At the stroke of a pen", Paulston et al. (2007: 386) write, members of the dominant ruling power can become minorities in a newly independent state. Extra and Gorter (2008) discuss the approach in the introduction to

their own framework for regional minority languages in Europe. They opt "for a simple typology" (Extra and Gorter 2008: 26) for their framework, which has five categories: (i) languages spoken in only one member state of the European Union; those spoken in more than one − whether (ii) unofficial in each country, or (iii) official in some; (iv) those three varieties (Lëtzebuergesch, Irish, and Maltese) that are 'small', but which have official status; and (v) non-territorial languages (notably, Romani and Yiddish). Tsunoda (2006: 49) uses the model as the scaffolding for his chapter on endangered languages ("we shall adopt Edwards's […] typology, which sets up eleven groups according to which various factors may be classified"). In so doing, he reminds us that typologies are a subsection of language ecology. Unfortunately, the great potential value of Tsunoda's monograph − as a comprehensive survey of endangered languages − is curtailed by a rambling and often indigestible presentation. As Mühlhäusler (2007: 105) notes, the confusion and lack of coherence here are particularly disappointing in a volume meant to be "a textbook or a guide for practitioners".

Knooihuizen (2006) notes that the basic emphasis of the model has been on minority-language situations, and views this as a flaw − but one which he implies could be easily taken care of. In fact, as readers can easily grasp, it was clear from the beginnings of the model that very slight alterations − suitable expansions of the cell questions, in particular − might make the framework applicable in non-minority settings, and thus a useful guide to the description of any language and any contact context. Its initial emphasis on minority-language matters emerged for specific purposes (as Table 45.1. indicates) − but its thrust was only briefly seen as restricted to 'small' varieties. Knooihuizen also points to what he sees as the synchronic aspect of the model, whereas it should always have been clear that no particular time restrictions need apply. He makes reference to Labov's observations about the 'bad data problem' that bedevils historical linguistics, which Labov (1994: 11) argues is generally best seen as "the art of making the best use of bad data", a reference to some of the obvious limitations of an incomplete and/or inaccurate historical record "that cannot be compensated for". Indeed, historical embedding should always be a central concern, because the value of social-scientific work that lacks contextualization is fleeting at best. Interestingly, Darquennes (2013) − see also below − removed 'history' from the Table 45.2 framework, making the reasonable argument that all the other perspectives can and should have an historical dimension or aspect. This is true enough, but perhaps retaining a specific category would highlight the need to avoid one-off snapshots. In any event, in an illuminating representation of that tabular framework, as applied in his specific case study (of Shetland Norn), Knooihuizen (2006) fills each of the cells with one of three colors: green, to represent areas in which good data exist; yellow, for reasonable or fair degrees of data; and red, for 'bad', i.e. inadequate or insufficient, data.

Darquennes (2013) presents a tabular representation of the 33-cell framework shown in this article, and importantly, offers some suggestions for the expansion of my typological scaffolding. He cites my own observations about the tentative and limited nature of the latter, but acknowledges its "conceptual clarity" (Darquennes 2013: 105). He further suggests some quite specific alterations: beyond his treatment of 'history' (noted in the previous paragraph), he replaces the 'speaker' category with 'group' and elaborates on the list of 'disciplinary perspectives'.

A final note here: Vail (2006) employed the model in his assessment of Northern Khmer, in which he notes that cultural and social anthropology are "curious omissions

from [my] otherwise comprehensive list" (Vail 2006: 144). He is right, of course − and there are no doubt many other fine-grained perspectives that could reasonably be included. I did think, however, that the sociological and linguistic perspectives would be sufficient, since their application would necessarily have anthropological − or anthropology-of-language − connotations. Overall, Vail (2006: 140) refers to my typology as "the most robust model" of both macro- and micro-level approaches to the ecology of endangered languages.

6. A concluding remark

Further developments in language profiling of the sort discussed here will obviously benefit from the applications and criticisms presented in the previous section. I am of course glad that the work has attracted some favorable notice, and that it has elicited suggestions for modifications; it is clear that much more needs to be done. Nonetheless, based upon the work of my predecessors − and recalling specifically the words of Haugen and Ferguson, cited in the first section − the exercise appears eminently worthwhile. Even a thoroughgoing multivariate checklist would be of service, while a comprehensive typology could be a useful tool for description and comparison, which in turn could lead to more complete conceptualizations of minority-language situations, be a heuristic for additional, more systematic investigations, and perhaps permit predictions concerning language shift and/or maintenance outcomes. Finally, the utility of the framework may be increased, given that it can easily be made applicable to cover languages and settings other than those involving 'small' or threatened varieties, and that the questions it prompts − provided there is a suitable historical underpinning − can free it from merely synchronic applications. I am grateful to Knooihuizen (2006) for reminding me to emphasize these expanded possibilities.

Taxonomic work should be seen in broad-brush terms as part of the larger sociolinguistic or sociology-of-language enterprise. At a more specific level, it finds an obvious niche in the ecology-of-language literature (see Edwards [2008, 2009, 2010] and Darquennes [2013] for further discussion), which is to say that it provides a perspective on the central issue of individual and group identity. So understood, typological effort reminds us that all relevant and enduring social-scientific work must rest upon a classificatory base. To formalize matters a bit is no bad thing, particularly in contexts in which the applications of research may have some immediate benefits, and in which ignorance of on-the-ground variables may lead to disappointment.

7. References

Ammon, Ulrich

2011 A checklist of sociolinguistic language maintenance indicators for diaspora minorities (with a focus on German examples). In Bruno Moretti, Elena Maria Pandolfi & Matteo Casoni (eds.), *Vitalità di una lingua minoritaria: aspetti e proposte metodologiche*, 43−63. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.

Anderson, Alan
1981 The problem of minority languages: Reflections on the Glasgow conference. *Language Problems and Language Planning* 5. 291–303.
Austin, Peter & Julia Sallabank (eds.)
2014 *Endangered languages: Beliefs and ideologies in language documentation and revitalization*. Oxford: Oxford University Press.
Berruto, Gaetano
2011 Considerazione conclusive. In Bruno Moretti, Elena Maria Pandolfi & Matteo Casoni (eds.), *Vitalità di una lingua minoritaria: Aspetti e proposte metodologiche*, 289–302. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.
Bourhis, Richard, Howard Giles & Doreen Rosenthal
1981 Notes on the construction of a 'Subjective Vitality Questionnaire' for ethnolinguistic groups. *Journal of Multilingual and Multicultural Development* 2. 145–155.
Clyne, Michael
2003 *Dynamics of language contact.* Cambridge: Cambridge University Press.
Darquennes, Jeroen
2013 The contribution of the ecology of language to the advancement of linguistic profiling: Some notes and some preliminary suggestions on further improvements. In Wim Vandenbussche, Ernst Håkon Jahr & Peter Trudgill (eds.), *Language ecology for the 21st century: Linguistic conflicts and social environments*, 95–114. Oslo: Novus.
Edwards, John
1991 Socio-educational issues concerning indigenous minority languages: Terminology and status. In Jantsje Sikma & Durk Gorter (eds.), *European lesser used languages in primary education*, 207–226. Ljouwert: Fryske Akademy/Mercator.
Edwards, John
1992 Sociopolitical aspects of language maintenance and loss: Towards a typology of minority language situations. In Willem Fase, Koen Jaspaert & Sjaak Kroon (eds.), *Maintenance and loss of minority languages*, 37–54. Amsterdam: John Benjamins.
Edwards, John
2008 The ecology of language: Insight and illusion. In Angela Creese, Peter Martin & Nancy Hornberger (eds.), *Encyclopedia of language and education*, vol. 9 (*Ecology of language*), 15–26. New York: Springer.
Edwards, John
2009 *Language and identity.* Cambridge: Cambridge University Press.
Edwards, John
2010 *Minority languages and group identity: Cases and categories.* Amsterdam: John Benjamins.
Ehala, Martin
2015 Ethnolinguistic vitality. In Karen Tracy, Cornelia Ilie & Todd Sandel (eds.), *International encyclopedia of language and social interaction*, 1–7. Oxford: Wiley.
Eisenberg, Avigail & Jeff Spinner-Halev (eds.)
2005 *Minorities within minorities.* Cambridge: Cambridge University Press.
ELDIA
2010–2013 EuLaViBar Project http://www.eldia-project.org/ (last accessed on 27 December 2018). [ELDIA = European Language Diversity for All; EuLaViBar = European Language Vitality Barometer].
ELDIA EuLaViBar Toolkit
2013 https://phaidra.univie.ac.at/view/o:301101/. (accessed on 13 June 2018).
Extra, Guus & Durk Gorter
2008 The constellation of languages in Europe: An inclusive approach. In Guus Extra & Durk Gorter (eds.), *Multilingual Europe: Facts and policies*, 3–60. Berlin: Mouton de Gruyter.

Ferguson, Charles
1962 The language factor in national development. In Frank Rice (ed.), *Study of the role of second languages in Asia, Africa, and Latin America*, 8–14. Washington: Center for Applied Linguistics.
Ferguson, Charles
1991 Diglossia revisited. *Southwest Journal of Linguistics* 10. 214–234.
Fishman, Joshua
1991 *Reversing language shift.* Clevedon: Multilingual Matters.
Gawrisch, Vanessa
2011 Methodische Überlegungen zur Messung der Spracherhaltchancen am Beispiel der deutschen und deutschsprachigen Minderheiten in der EU. In Bruno Moretti, Elena Maria Pandolfi & Matteo Casoni (eds.), *Vitalità di una lingua minoritaria: aspetti e proposte metodologiche*, 227–236. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.
Giles, Howard, Richard Bourhis & Donald Taylor
1977 Towards a theory of language in ethnic group relations. In Howard Giles (ed.), *Language, ethnicity and intergroup relations*, 307–348. London: Academic Press.
Grenoble, Lenore & Lindsay Whaley (eds.)
1998a *Endangered languages: Current issues and future prospects*. Cambridge: Cambridge University Press.
Grenoble, Lenore & Lindsay Whaley
1998b Toward a typology of language endangerment. In Lenore Grenoble & Lindsay Whaley (eds.), *Endangered languages: Current issues and future prospects*, 22–54. Cambridge: Cambridge University Press.
Grenoble, Lenore & Lindsay Whaley
2006 *Saving languages: An introduction to language revitalization.* Cambridge: Cambridge University Press.
Haarmann, Harald
1986 *Language in ethnicity: A view of basic ecological relations.* Berlin: Mouton de Gruyter.
Haarmann, Harald
1996 Ökolinguistik. In Hans Goebl, Peter Nelde, Zdeněk Starý & Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact linguistics / Linguistique de contact*, 842–852. Berlin: De Gruyter.
Haugen, Einar
1972 The ecology of language. In Anwar Dil (ed.), *The ecology of language: Essays by Einar Haugen*, 325–339. Stanford: Stanford University Press.
Hyltenstam, Kenneth & Christopher Stroud
1996 Language maintenance. In Hans Goebl, Peter Nelde, Zdeněk Starý & Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact linguistics / Linguistique de contact*, 567–578. Berlin: De Gruyter.
Iannàccaro, Gabriele & Vittorio Dell'Aquila
2011 Numeri soggettivi: Spunti sulla vitalità linguistica da inchieste e trattamenti quantitativi. In Bruno Moretti, Elena Maria Pandolfi & Matteo Casoni (eds.), *Vitalità di una lingua minoritaria: Aspetti e proposte metodologiche*, 151–192. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.
Jones, Mari (ed.)
2015 *Policy and planning for endangered languages*. Cambridge: Cambridge University Press.
King, Kendall
2001 *Language revitalization processes and prospects: Quichua in the Ecuadorian Andes.* Clevedon: Multilingual Matters.

Kloss, Heinz
1967 Types of multilingual communities: A discussion of ten variables. In Stanley Lieberson (ed.), *Explorations in sociolinguistics*, 7–17. Bloomington: Indiana University Press.
Knooihuizen, Remco
2006 Language shift, ethnolinguistic vitality and historical sociolinguistics: Testing the models. Paper presented at the Linguistics and English Language Postgraduate Conference, University of Edinburgh, 20–21 April. http://www.lel.ed.ac.uk/~pgc/archive/2006/2006knooihuizen.pdf/. (accessed on 13 June 2018).
Krauss, Michael
1992 The world's languages in crisis. *Language* 68. 4–10.
Labov, William
1994 *Principles of linguistic change*, vol. 1 (*Internal factors*). Oxford: Wiley-Blackwell.
Lewis, Paul
2005 *Towards a categorization of endangerment of the world's languages.* Dallas: Summer Institute of Linguistics.
Moretti, Bruno & Elena Maria Pandolfi
2011 Outlining a vitality index for Italian in Switzerland. In Emanuele Miola & Paolo Ramat (eds.), *Language contact and language decay: Socio-political and linguistic perspectives*, 191–211. Pavia: Istituto Universitario di Studi Superiori.
Moretti, Bruno, Elena Maria Pandolfi & Matteo Casoni
2011a Introduzione / Introduction. In Bruno Moretti, Elena Maria Pandolfi & Matteo Casoni (eds.), *Vitalità di una lingua minoritaria: Aspetti e proposte metodologiche*, 7–22. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.
Moretti, Bruno, Elena Maria Pandolfi & Matteo Casoni (eds.)
2011b *Vitalità di una lingua minoritaria: Aspetti e proposte metodologiche*. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.
Mufwene, Salikoko
2005 Globalization and the myth of killer languages. In Graham Huggan & Stephan Klasen (eds.), *Perspectives on endangerment*, 19–48. Hildesheim: Olms.
Mühlhäusler, Peter
2007 Review of *Language endangerment and language revitalization* (Tasaku Tsunoda). *Current Issues in Language Planning* 8. 102–105.
Paulston, Christina, Szidonia Haragos, Verónica Lifrieri & Wendy Martelle
2007 Some thoughts on extrinsic linguistic minorities. *Journal of Multilingual and Multicultural Development* 28. 385–399.
Russell, Bertrand
1950 The superior virtue of the oppressed. In Bertrand Russell, *Unpopular Essays*, 80–87. London: Allen & Unwin.
Smith, Benjamin, Martin Ehala & Howard Giles
2017 Vitality theory. In Jon Nussbaum (ed.), *Oxford research encyclopedia of communication*, 1–22. Oxford: Oxford University Press.
Stähli, Adrian
2011 Aspetti di vitalità, mantenimento e perdita di una lingua: Riflessioni per un inquadramento sociolinguistico di Bosco Gorin, comune walser in Ticino. In Bruno Moretti, Elena Maria Pandolfi & Matteo Casoni (eds.), *Vitalità di una lingua minoritaria: Aspetti e proposte metodologiche*, 211–226. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.
Stewart, William
1962 An outline of linguistic typology for describing multilingualism. In Frank Rice (ed.), *Study of the role of second languages in Asia, Africa, and Latin America*, 15–25. Washington: Center for Applied Linguistics.

Stewart, William
1968 A sociolinguistic typology for describing national multilingualism. In Joshua Fishman (ed.), *Readings in the sociology of language*, 531−545. The Hague: Mouton.
Tsunoda, Tasaku
2006 *Language endangerment and language revitalization.* Berlin: Mouton de Gruyter.
UNESCO Ad Hoc Expert Group on Endangered Languages
2003 *Language vitality and endangerment.* Paris: UNESCO.
UNESCO
2010 *Atlas of the world's languages in danger*, 3rd edn. Paris: UNESCO. http://www.unesco.org/new/en/culture/themes/endangered-languages/atlas-of-languages-in-danger/. (accessed on 13 June 2018).
Vail, Peter
2006 Can a language of a million speakers be endangered? Language shift and apathy among Northern Khmer speakers in Thailand. *International Journal of the Sociology of Language* 178. 135−147.
Valiquette, Hilaire
1998 Community, professionals, and language preservation. In Nicholas Ostler (ed.), *Endangered languages*, 107−112. Bath: Foundation for Endangered Languages.
White, Paul
1987 Geographical aspects of minority language situations in Italy. Paper presented at the International Seminar of Geolinguistics, Staffordshire Polytechnic, 24−26 May.
Yağmur, Kutlay & Sjaak Kroon
2003 Ethnolinguistic vitality perceptions and language revitalisation in Bashkortostan. *Journal of Multilingual and Multicultural Development* 24. 319−336.
Yağmur, Kutlay & Sjaak Kroon
2006 Objective and subjective data on Altai and Kazakh ethnolinguistic vitality in the Russian Federation Republic of Altai. *Journal of Multilingual and Multicultural Development* 2. 241−258.

John Edwards, Halifax (Canada)

46. Social network analysis

1. Introduction
2. Approaches to social networks
3. Applications of SNA in language research
4. SNA in language contact research
5. SNA and bilingual acquisition
6. Conclusion
7. References

1. Introduction

Social network analysis (SNA), in a broad sense, is the study of connections that exist between individuals, groups, organizations, and other entities in an ongoing system of exchange. When it is applied to larger, institutionalized systems in order to understand

https://doi.org/10.1515/9783110435351-046

management and decision-making structures, it is often known as organizational network analysis. When applied to individuals and social groups, SNA is primarily concerned with mapping and measuring the location and grouping of members of society, their relationships, and the flow between them. These measures give us insights into the various roles individuals play in a network (i.e. the connectors, mavens, leaders, bridges, isolates, etc.) as well as the structural patterns of the network as a whole (i.e. dense or loose clusters, pathways, the center, and the periphery). The roles within, and structural patterns of, the networks can then be used to investigate different behaviors of individual members of society. SNA is essentially a quantitative analytical measure that can provide a visual or computational representation of the connections between individuals and of the overall structure of a network. The popular notion of a social network, i.e. social media, has nothing to do with SNA, although SNA can clearly be applied to analysis of social media use.

SNA has been applied to the study of language contact and contact-induced linguistic change to explicate the social mechanisms underlying variations in the linguistic behaviors of speakers, addressing questions such as: What effect do our relationships with others have on our language practices? What role does language play in our developing social relations? How do changes in group membership affect individual members' behavior? Are our attitudes towards others and their languages related to our relationships with them and our social group memberships? SNA can in principle be applied to studying entire communities, for example, to compare the effects of different network structures on language maintenance and language shift.

2. Approaches to social networks

A social network is defined as the web-like pattern of relationships between individuals (Gelles and Levine 1999: 207). SNA as a major analytical approach has two different, but related, origins (see Scott [2012] for details): one in the Chicago and later in the Harvard schools of structural sociology and social psychology, particularly sociometry (as exemplified in the work of White, Wellman, Berkowitz, and Granovetter), and the other in the Manchester school of cultural anthropology (associated with the work of Barnes, Mitchell, and Bott). Two broad approaches to the description and analysis of the relationship patterns can be distinguished: the whole network (sociological) approach and the egocentric network (anthropological) approach. Both approaches focus on the quantitative measurement of relationships between people, with the former focusing on the overall structure of the network and the latter on the profile (size and composition) of the ties. For example, the Harvard group's main contribution was the development of mathematical models of network analysis, in which concepts such as equivalence and centrality are key, whereas the Manchester group was more interested in the effect of integration into a group on individuals' behaviors and attitudes. The basic assumptions are that the structural pattern of one's social network, and the specific role one has within that network, impact one's everyday behavior(s). SNA is now applied to a wide range of research fields, including psychology, political science, ecological and environmental sciences, public health, business management, information sciences, education, and linguistics.

It is important to note that social network analysts have a specific conceptualization of society. In their view, there is no 'social structure' as an independent entity. Society is instead viewed as being constituted of a series of social relationships. These relationships have a number of dimensions: time, space, and power. We can therefore talk about, for example, historical, geographical, and hierarchical networks. Social change is then also change in social relationships and networks. Moreover, social networks not only connect us with each other, but also provide an individual's social locale, which plays an important part in determining the specific information to which the individual is exposed. Individuals in different social locales may receive different information, and participate in different social events, and as a result, they may come to construct knowledge differently. Because the locales which social networks provide are closely related to patterns of social interaction(s), they also influence individuals' social expectations, which may in turn lead to certain types of sociocognitive bias(es). Social networks are therefore an important part of social cognition, because cognitive processes may directly involve individuals' perceptions of their social locale. Those at the center of social activities and social relationships receive and perceive information quite differently from people who are on the periphery.

As seen in Scott and Carrington (2011), SNA can be used to address issues such as:

(i) how communities and societies are structured, group boundaries, and the international dynamics of social groupings;
(ii) individual roles within communities and their relationship to the transmission of information, innovation, and so on;
(iii) the relationship between community structure, individual roles within the community, and individuals' orientation to means of production;
(iv) life cycle stages of individual members of a social group, and their relationship to the social position of the individual in the group and to network connectivity;
(v) the roles of information and communication technologies in structuring social relationships;
(vi) and the dynamic processes organizing networks, and their structural consequences.

While all of these are potentially relevant to the study of language contact, I focus in the following sections only on those aspects that have been systematically discussed and applied in language research.

3. Applications of SNA in language research

The concept of 'social network' has been used most extensively in sociolinguistics (especially in variationist sociolinguistics) both as a methodological approach to data collection, and as a theoretical framework for data analysis. Milroy and Gordon (2005), for example, describe the advantages of using a 'friend-of-a-friend' approach to access speakers, speech communities, and speech events which might otherwise have been difficult to gain access to. There are a number of advantages to such a network approach. First, it may help researchers to gain access to aspects of social life that may not usually be open to outsiders. For instance, it enabled Lesley Milroy to go into neighborhoods with different religious affiliations in Belfast, Northern Ireland, in the 1970s. Second, it offers researchers a specific perspective that is embedded in human social relations.

Suzanne Moffatt (1990), a white, English-speaking woman, was able to study the bilingual behavior of Pakistani children and their parents through the social networks she developed in her role as a teaching assistant in a school mainly attended by Pakistani children in Newcastle's West End, England. This perspective helped her to observe and understand the children's and parents' attitudes towards bilingualism. Third, it may help researchers contact a large number of people in a relatively short period of time.

In data analysis, social networks were initially used as a metaphor to describe speech communities which were either 'close-knit' or 'open'. In his studies of language variation in India, Gumperz (1971) found that informal friendship contacts were one of the key social factors accounting for phonetic variations in the dialects of Khalapur, a northern village in India. The phonetic features in the speech by members of close-knit networks were more 'focused', whereas those of open networks were 'diffused' (see also Gumperz 1982). Yet it was variationist sociolinguists who first used the social network as an analytic concept and specifically collected network data, which were then used in their explanation of linguistic variation and change.

Labov's (1972) study of Black teenagers in Harlem, New York City, investigated the linguistic behavior of members of street gangs known as the Thunderbirds and the Aces, as well as that of outsiders in the same age group, known as the Lames. Adolescents worldwide are notorious for their tendency to form dense networks and place great value on being 'part of the gang'; these dense networks exert peer pressure on members' behavior, including their use of language. Labov discovered that the gang members in the Thunderbirds and the Aces showed much higher frequencies of nonstandard vernacular variants. The Lames, on the other hand, used such vernacular features less frequently. Similarly, Eckert (2000) in her study of Detroit teenagers' communities of practice identified two networks, called the Jocks and the Burnouts, that differed essentially in the role of high school culture in their mundane activities: the Jocks were into sports, school politics, and other social and academic activities, whereas the Burnouts were outsiders, hung out in restaurants and malls, and got themselves into trouble from time to time. Eckert found significant linguistic correlations depending on network affiliations: the Jocks displayed more middle class variants, while the Burnouts used more typically working class variants.

Both Labov's (1972) and Eckert's (2000) studies adopted the whole network approach, and used social networks to define social differentiation in language with reference to social class and ethnic segmentation of society. More recent sociolinguistic studies using the whole network approach include: Tribur (2017), who examined the role of network structures in dialectal diversity in Amdo Tibetan; Bergs (2005), who examined morphosyntacitc variation in the Paston Letters (one of the largest collections of 15th century English private correspondence) from a historical-sociolinguistic perspective; Lev-Ari (2016), who analyzed the effect of network size on lexical access from a psycholinguistic perspective; and Fagyal et al. (2010), who developed a computational model to investigate the emergence of norms in simulated networks of linguistic influence. There are also applications of SNA in the study of online, digital, and technology mediated communication (e.g. Bergs 2006; Paolillo 1999).

Perhaps the best known sociolinguistic application of SNA is Milroy's (1987) study of phonological variation and change in three working class communities in Belfast, Northern Ireland. Instead of looking at whole network structures, Milroy focused on individuals' network ties, and the roles of such ties in speakers' use of phonological

variants. This approach to SNA is different from the whole network approach: it focuses on the individual, uses them as the anchor person, and observes the network contacts they have for different social purposes. Milroy (1987) developed a Network Strength Scale, which quantifies speakers' network characteristics with reference to various relationships of kin, work, and friendship. Speakers scored one point for each of the following conditions they satisfied:

- were members of a high-density, territorially based group (e.g. a bingo or card-playing group, a gang or a football team, or football supporters' club);
- had kinship ties with more than two households in the neighborhood;
- worked in the same place as at least two others from the neighborhood;
- worked in the same place as at least two others of the same sex from the neighborhood;
- associated voluntarily with workmates in leisure hours (Milroy and Gordon 2003: 121).

A series of statistical analyses revealed that speakers with the strongest vernacular features were generally those whose neighborhood network ties were the strongest.

As Milroy and Gordon (2003: 120) point out, a social network approach is attractive to variationist sociolinguists for a number of reasons. First, it provides a set of procedures for studying small groups wherein speakers are not discriminable in terms of any kind of social class index. This is, however, not to say that social networks are not linked to social class. In fact, there is a lot of evidence that social network patterns interact with socioeconomic class and gender. Milroy and Milroy (1992), for example, used Højrup's (1983) notion of 'life-mode' to argue that a dynamic model of class as a process, which splits the community into subgroups characterized by different network structures and different orientations to work, leisure, and family, is helpful in constructing an integrated theory of variation and change. A second advantage is that, since 'social network' is intrinsically a concept which relates to local practices, it has the potential to elucidate social dynamics driving language variation and change. Social networks develop and change over time, and are often associated with a range of social and psychological factors (Kossinets and Watts 2006; Snijders 2011; Suitor, Wellman, and Morgan 1997); in this regard, studies of language maintenance and shift in bilingual communities have made important contributions (see additional studies below). Finally, network analysis of the ego-centered kind, as exemplified in Milroy's (1987) study, offers a procedure for dealing with variation between individual speakers, as opposed to between groups constructed with reference to predetermined social categories.

One of the pressing social issues that social network analysts have been dealing with in the past decade or so is migration. As Ryan and D'Angelo (2018) and Wissink and Mazzucto (2018) demonstrate, SNA offers a practical and insightful tool for studying migration networks and the subsequent integration of migrants into their new place of settlement. The linguistic consequences of migration have also emerged as a topic in sociolinguistics, as exemplified in the studies in a special issue of the *Journal of Sociolinguistics* edited by Milroy (2002). Earlier applications of SNA by sociolinguists relied on Granovetter's (1973) notion of 'the strength of weak ties', focusing largely on how newly arriving immigrants access majority language norms and how certain linguistic features became diffused (see studies in Milroy [2000]). More recent studies have paid much more attention to how bilingual and multilingual norms are developed (Cheshire et al. 2011; Sharma 2017).

4. SNA in language contact research

The capacity of SNA to compare individual differences in the degree of integration into relationships that exert normative pressures on individuals' behaviors is particularly attractive to the study of immigrant and transnational communities. Such communities typically have to deal with issues of social integration during the course of migration and resettlement, and success in their integration efforts has long-term consequences for development of the community in terms of identity, internal structure, and relationships with other groups. Among the earliest studies of social network structures, integration, and language use in bilingual communities was Gumperz' (1982) account of the Slovenian-German community in a remote part of Austria's Gail Valley. He found that members of the poor, socially stigmatized farming community had traditionally been embedded in close-knit networks of mutual support. However, such behaviors changed as the economy shifted from dependence on subsistence farming to a service economy. Improved transportation and communication systems also meant that a host of other changes affected network structures and everyday social practices, including language practices. As interactions with urban outsiders increased, reliance on the local support network diminished. Use of German also increased significantly over time, while use of Slovenian alone was confined to specific traditional networks.

Gal's (1979) study of language shift in Oberwart, a village on the Austrian-Hungarian border, found similar patterns. German-Hungarian bilingualism has existed in Oberwart since before the 16th century. However, the two languages were functionally quite different: German was a language to be used with outsiders, and Hungarian was the linguistic symbol of group identity for the Oberwart peasants. This state of affairs began to change after the Second World War, when local the industrial and commercial economy developed so significantly that nonagricultural employment was an increasingly attractive option. Gal found a parallel language shift from Hungarian-dominant bilingualism to German-dominant bilingualism. She ranked individual speakers according to their degree of 'urbanization' or 'Austrianness', and observed a strong correlation between their language choice patterns and social network contacts.

In his study of language choice and language shift in three generation Chinese families in the UK, Li Wei (1994) specifically collected social network data, which were analyzed in terms of ethnic and peer group composition in three different types of networks: exchange, interactive, and passive networks. Exchange networks are collectives of people with whom the probability of rewarding exchanges (or unrewarding exchanges, in cases of conflict-habituated relationships) is high (Milardo 1988). In other words, these are the people with whom the ego not only interacts routinely, but also exchanges direct aid, advice, and support, as well as criticism and interference. Although in principle an individual's social networks can be infinite, empirical evidence suggests that, for practical reasons, exchange networks are effectively limited to first order contacts, which normally number between 20 and 30 people (Mitchell 1986). Interactive networks, on the other hand, consist of people with whom the ego interacts frequently, and perhaps over prolonged periods of time, but crucially, the probability of rewarding exchanges is low, that is, the ego does not rely on these contacts for personal favors and other material (or symbolic) resources. An example of interactive ties would be a shop owner and his or her customers. Finally, the identification of 'passive networks' was particularly useful for studying migrant communities. Passive ties are marked by an absence of regular con-

tact, but are equally considered important by the ego, who depends on such relationships for sentimental and moral support or for influence. Many migrants, for example, would have relatives and friends who, for various reasons, are physically distant from them, but who are still regarded as important relations. Migrants often cling psychologically to their traditional ties, despite the passage of time. Passive ties can be 'activated' through home visits and other interactions.

Adopting an ego-network approach, Li Wei collected information on the social contacts of 58 individuals in the Chinese community. For the exchange networks, participant observation elicited an initial list of around 30 contacts per individual. The list was then presented to the speaker for verification, and revised accordingly. A resulting 20 non-kin contacts were used as the basis for analyzing individual speakers' exchange networks. In order to examine individuals' degree of social integration, two network indices were constructed: an ethnic index, which was calculated in terms of the number of Chinese versus non-Chinese ties out of the 20 exchange contacts for each speaker, and a peer index, reflecting the number of people belonging to the same generation as the speaker (as opposed to those belonging to other generations, whether older or younger). These indices were used to test two basic hypotheses (see further Li Wei 1994: 121–122):

- Speakers whose exchange networks consist of a relatively large number of ethnic (Chinese) ties would display more traditional social behavior, such as using Chinese-dominant language choice patterns, while those with fewer ethnic ties within their exchange networks would have moved away from such traditional patterns and adopted English-oriented behavior.
- Speakers whose exchange networks consist of a relatively large number of peer ties would display behavior which conforms to the overall pattern of the generation to which they belong.

Notice here that the indicators used to measure personal networks are different from those used by Gal or Milroy. Milroy (1987: 141) comments on the principles for selecting appropriate network indictors and designing network measures. They must first reflect conditions which have repeatedly been found to be important in a wide range of network studies in predicting the extent to which normative pressures are applied by a local community; and second, they must be recoverable from data collected in the field and be easily verifiable.

In the same vein, Li Wei (1994) constructed 'ethnic' and 'peer' indices for interactive and passive ties. However, whereas the numbers for the exchange networks indicated specific individuals, those for interactive networks represented proportions. Each speaker was observed to determine how many people there were with whom they routinely interacted without exchange of material and/or moral support; the total number of interactive ties obviously varied. In addition, ten passive ties were analyzed for each speaker in the sample. The percentage of ethnic and peer contacts were then calculated.

Using Analysis of Variance (ANOVA), Li Wei (1994) found that the ethnic index of exchange networks was the best predictor for an individual's language choice patterns: the more Chinese contacts a speaker had in the exchange network, the more likely the speaker was to speak Chinese all the time. The ethnic index also showed a similar pattern in interactive networks, albeit to a lesser extent. However, peer contacts worked in very complex ways: the peer composition of exchange networks contributed to language maintenance *within* generations on the one hand, and to language shift (i.e. from Chinese

to English) *across* generations on the other. This was even more true of interactive networks: the more peer-group contacts (members of the same generation) a grandparent had, the more Chinese was used; the more contacts a parent had with members of the grandparent generation, the more Chinese was used, while the more contacts a parent had with members of the child generation, the more bilingual s/he was; and the more peer-group contacts a child had, the more English was used.

This kind of social network analysis also allowed Li Wei (1994) to compare individual differences in language choice patterns. Individuals' social networks can be very different in terms of composition and size. The sociolinguistic hypothesis is that behavioral differences in language practices between individuals are conditioned by the networks in which they operate. In Li Wei's study, he revealed that two of the male grandparents, for example, used a lot more English than other members of their own generation. In fact, their language choice patterns were more typical of the parent generation. When social network contacts were examined, these two individuals were shown to have significantly more non-Chinese contacts in both their exchange and interactive networks than did their peers.

In a similar vein, Lanza and Svendsen (2007) analyzed social network contacts of members of the Filipino community in Oslo, Norway, with regard to maintenance of the ethnic language versus shift to Norwegian and English. They argued that individuals' social network composition could be a useful predictor of language choice, language maintenance, and language shift. Other studies of language choice, language maintenance, and language shift in migrant communities that used SNA include: Bortoni-Ricardo's (1985) work on rural dialect speakers in Brazil who migrated to a satellite city of the capital Brasilia, where she found that the social network ties of migrant workers, especially whether or not they were kinsfolk, or whether said ties had been contracted in the premigration period, were closely correlated with use of standard language versus dialect features; Schooling's (1990) work on language maintenance in Melanesia, where he developed a typology of social networks that corresponded to observed patterns of language choice; Stoessel's (2002) study of immigrant women in the US, where 'maintainers' and 'shifters' were differentiated by both language choice patterns and social network contacts; Hulsen, de Bot, and Weltens' (2002) study of language shift in three generations of Dutch migrants in New Zealand; Matsumoto's (2010) study of the postcolonial multilingual island of Palau; and Sallabank's (2010) study of the role of social networks in endangered language maintenance (and revitalization) of Guernesiais in the Channel Islands.

5. SNA and bilingual acquisition

Although sociolinguists who use the concept of social network in their work take a particular interest in changes over time, the approach they use tends to be cross-sectional rather than longitudinal, i.e. they focus on different groups of speakers at a specific time, rather than following one group of speakers over time. Yet social networks are developed over time, and are subject to constant changes. Applications of social network analysis in other fields have looked at the time dimension of network development and change, and its impact on individuals' behavior(s) (e.g. Cochran and Brassard 1979).

Li Wei (2011) specifically addresses the issue of change in social network ties in a longitudinal study of three Chinese children (two girls and one boy) who were first exposed to English at around one year of age, and were developing as Chinese-English bilinguals. Using a combination of research instruments and data sources, including parental checklists (a CDI-style Chinese vocabulary checklist, and a British English version of the MacArthur-Bates CDI), weekly recordings of the family engaged in joint activities, parental diaries (language choice, social activities, and developmental features and milestones), monthly recorded home observations, and interviews with parents, Li Wei was able to track changes in the quantity and quality of the linguistic input the children received, and the effects thereof on lexical development (measured in terms of lexical diversity) and development of code-switching over time. As in his earlier work on language maintenance and language shift in the Chinese community in the UK, Li Wei focused on the composition of social networks (e.g. peer versus non-peer, Chinese versus non-Chinese, bilingual versus monolingual), this time including the size of the network (in particular, its growth over time). The following is a list of highlights from his findings *vis-à-vis* the social network effect:

(i) Sibling effect
A sibling forms an important part of a child's social network; both second-born subjects reached the two-word stage 3–5 months sooner than the singleton.

(ii) Interlocutor effect
   a) There was a strong correlation between the first language of the primary caregiver (i.e. the person who spends the most time engaged in joint activities with the child), and the child's language dominance (measured in terms of language choice preference and lexical diversity).
   b) Code-switching was restricted to specific interlocutors (i.e. bilingual parents).

(iii) Changes over time
Language development over time is strongly correlated with (changes in) the primary caregiver's language dominance or preference.
   a) Lexical diversity in the two languages combined grew according to the number of contacts the child had *outside* the family.
   b) Attendance at an English-language nursery led to reduced code-switching (due to reduced time engaged in joint activities with parents).
   c) As the child grew older, the two languages developed in very separate contexts. The type of joint activities in which children engaged differed in different contexts, and code-switching patterns changed (increased intersentential code-switching, with intrasentential code-switching restricted to English content words.

Li Wei's (2011) study demonstrates the potential of social network analysis as a practical, quantitative method of measuring variation and change in language input, and the effect on the language development of young children. It also helps break the artificial boundaries between sociolinguistics and developmental psycholinguistics with regard to language socialization. A similar study is Velázquez (2013), who examined mothers' social networks and children's home language maintenance. An interesting extension of this work is found in recent studies on the effect of study abroad on second language learning, where the structure as well as composition of personal network ties are found to have a significant impact on students' language learning (Dewey, Belnap, and Hillstrom 2013; Mitchell, Tracy-Ventura, and McManus 2017).

6. Conclusion

Social networking is a pervasive fact of human life. Social networks are as indispensable as language to members of any society, and provide an important social locale for their members. An individual's social locale plays a critical part in determining the specific information to which that individual is exposed. People in different social locales receive information in different ways, and participate in different social events; as such, they may come to construct knowledge differently. Social networks therefore influence individuals' social expectations, which may lead to certain types of sociocognitive bias(es), which in turn may affect how they construct knowledge (Scott and Carrington 2011).

SNA has been proven to be a very useful avenue for understanding the linguistic dimensions of human life. It is analytically much more useful than other metaphorical notions, such as speech community and community of practice, as it offers a concrete, practical way to analyze social structures and individuals' positions therein. It has been integrated into the study of language to illuminate synchronic variation and diachronic change, and has also been applied to the study of bilingualism and multilingualism, providing findings and theories which specifically account for phenomena such as language choice, language maintenance, and language shift. More recently, SNA has been used to understand the social conditions of child language development, second language learning, and technology-mediated communication. More work can be done on lifespan changes in language and social networks, as well as on the impact new social networking media have on language practice and language structures.

Acknowledgments

This chapter is a substantially reworked version of Daming, Xiamei, and Li Wei (2008). The author wishes to thank his former co-authors, as well as Wiley-Blackwell, for their permission to reuse part of the chapter.

7. References

Bergs, Alexander
2005 *Social networks and historical sociolinguistics: Studies in morphosyntactic variation in the Paston Letters*. Berlin: Walter de Gruyter.

Bergs, Alexander
2006 Analyzing online communication from a social network point of view: Questions, problems, perspectives. *Language@Internet* 3(3), http://www.languageatinternet.org/articles/2006/371. (accessed on 10 September 2018).

Bortoni-Ricardo, Stella Maris
1985 *The urbanization of rural dialect speakers: A sociolinguistic study in Brasil*. Cambridge: Cambridge University Press.

Cheshire, Jenny, Paul Kerswill, Sue Fox & Eivind Torgersen
2011 Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics* 15(2). 151–196.

Cochran, Moncrieff & Jane Anthony Brassard
1979 Child development and personal social networks. *Child Development* 50(3). 601–616.

Daming, Xu, Wang Xiaomei & Li Wei
2008 Social network analysis. In Li Wei & Melissa G. Moyer (eds.), *The Blackwell guide to research methods in bilingualism and multilingualism*, 263–274. Oxford: Blackwell.
Dewey, Dan. P., R. Kirk Belnap & Rebecca Hillstrom
2013 Social network development, language use, and language acquisition during study abroad: Arabic language learners' perspectives. *Frontiers: The Interdisciplinary Journal of Study Abroad* 22. 84–110.
Eckert, Penelope
2000 *Language variation as social practice: The linguistic construction of identity in Belten High*. Oxford: Wiley-Blackwell.
Fagyal, Zsuzsanna, Aamarth Swarup, Anna María Escobar, Les Gasser & Kiran Lakkaraju
2010 Centers and peripheries: Network roles in language change. *Lingua* 120. 2061–2079.
Gal, Susan
1979 *Language shift: Social determinants of linguistic change in bilingual Austria.* San Francisco: Academic Press.
Gelles, Richard J. & Anne Levine
1999 *Sociology: An introduction*. Boston: McGraw-Hill.
Granovetter, Mark S.
1973 The strength of weak ties. *American Journal of Sociology* 78(6). 1360–1380.
Gumperz, John J.
1971 Dialect difference and social stratification in a north Indian village. In Anwar S. Dil (ed.), *Language in social groups: Essays by John J. Gumperz*, 25–47. Stanford: Stanford University Press.
Gumperz, John J.
1982 *Language and social identity.* Cambridge: Cambridge University Press.
Højrup, Thomas
1983 The concept of life-mode: A form-specifying mode of analysis applied to contemporary Western Europe. *Ethnologia Scandinavica* 13. 15–50.
Hulsen, Madeleine, Kees de Bot & Bert Weltens
2002 Between two worlds: Social networks, language shift, and language processing in three generations of Dutch migrants in New Zealand. *International Journal of the Sociology of Language* 153. 27–52.
Kossinets, Gueorgi & Duncan J. Watts
2006 Empirical analysis of an evolving social network. *Science* 311(5757). 88–90.
Labov, William
1972 *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.
Lanza, Elizabeth & Bente Ailin Svendsen
2007 Tell me who your friends are and I might be able to tell you what language(s) you speak: Social network analysis, multilingualism, and identity. *International Journal of Bilingualism* 11(3). 275–300.
Lev-Ari, Shiri
2016 How the size of our social network influences our semantic skills. *Cognitive Science* 40(8). 2050–2064.
Li Wei
1994 *Three generations, two languages, one family: Language choice and language shift in a Chinese community in Britain.* Clevedon: Multilingual Matters.
Li Wei
2011 The early acquisition of English as a second language: The case of young Chinese learners of English in Britain. In Annick De Houwer & Antje Wilton (eds.), *English in Europe today: Sociocultural and educational perspectives*, 95–112. Amsterdam: John Benjamins.

Matsumoto, Kazuko
2010 The role of social networks in the post-colonial multilingual island of Palau: Mechanisms of language maintenance and shift. *Multilingua* 29(2). 133–165.
Milardo, Robert M. (ed.)
1988 *Families and Social Networks*. Newbury Park, CA: SAGE.
Milroy, Lesley
1987 *Language and social networks*, 2nd edn. Oxford: Blackwell.
Milroy, Lesley
2000 Social network analysis and language change: Introduction. *European Journal of English Studies* 4(3). 217–223.
Milroy, Lesley
2002 Introduction: Mobility, contact and language change – Working with contemporary speech communities. *Journal of Sociolinguistics* 6(1). 3–15.
Milroy, Lesley & James Milroy
1992 Social network and social class: Toward an integrated sociolinguistic model. *Language in society* 21(1). 1–26.
Milroy, Lesley & Matthew Gordon
2003 *Sociolinguistics: Method and interpretation*. Oxford: Wiley Blackwell.
Mitchell, Clyde
1986 Network procedures. In Dieter Frick, Hans W. Hoefert, Heiner Legewie, Rainer Mackensen & Rainer K. Silbereisen (eds.), *The quality of urban life*, 73–92. Berlin: de Gruyter.
Mitchell, Rosamond, Nicole Tracy-Ventura & Kevin McManus
2017 *Anglophone students abroad: Identity, social relationships, and language learning*. London: Routledge.
Moffatt, Suzanne
1990 *Becoming bilingual: A sociolinguistic study of the communication of young mother tongue Panjabi-speaking children*. Newcastle upon Tyne, UK: University of Newcastle upon Tyne dissertation.
Paolillo, John
1999 The virtual speech community: Social network and language variation on IRC. *Journal of Computer-Mediated Communication* 4. http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1999.tb00109.x/full. (accessed on 10 September 2018).
Ryan, Louise & Alessio D'Angelo
2018 Changing times: Migrants' social network analysis and the challenges of longitudinal research. *Social Networks* 53. 148–158.
Sallabank, Julia
2010 The role of social networks in endangered language maintenance and revitalization: The case of Guernesiais in the Channel Islands. *Anthropological Linguistics* 52(2). 184–205.
Schooling, Stephen J.
1990 *Language maintenance in Melanesia: Sociolinguistics and social networks in New Caledonia*. Dallas: Summer Institute of Linguistics.
Scott, John
2012 *Social network analysis*, 3rd edn. London: SAGE.
Scott, John & Peter J. Carrington (eds.)
2011 *The SAGE handbook of social network analysis*. London: SAGE.
Sharma, Devyani
2017 Scalar effects of social networks on language variation. *Language Variation and Change* 29. 393–418.
Snijders, Tom A. B.
2011 Network dynamics. In John Scott & Peter J. Carrington (eds.), *The SAGE handbook of social network analysis*, 501–513. London: SAGE.

Stoessel, Saskia
2002 Investigating the role of social networks in language maintenance and shift. *International Journal of the Sociology of Language* 153. 93–132.
Suitor, Jill, Barry Wellman & David Morgan
1997 It's about time: How, why, and when networks change. *Social Networks* 19(1). 1–7.
Tribur, Zoe
2017 Social network structure and language change in Amdo Tibetan. *International Journal of the Sociology of Language* 245. 169–206.
Velázquez, Isabel
2013 Mother's social network and family language maintenance. *Journal of Multilingual and Multicultural Development* 34(2). 189–202.
Wissink, Marieke & Valentina Mazzucato
2018 In transit: Changing social networks of sub-Saharan African migrants in Turkey and Greece. *Social Networks* 53. 30–41.

Li Wei, London (UK)

47. Nexus analysis and multilingualism

1. Introduction
2. Nexus analysis: doing mediated discourse analysis
3. Nexus of practice
4. Interactional order, historical bodies, and circulating discourses: three main elements of nexus analysis
5. Activities of nexus analysis: Engaging, navigating and changing the nexus of practice
6. Language contact: a multilingual and multimodal turn
7. Conclusion: what type of research is nexus analysis suited for?
8. References

1. Introduction

Nexus analysis originated from a long-term research engagement in a multilingual setting. In *Nexus analysis and the emerging internet*, Scollon and Scollon (2004) describe and analyze how the use of new computer technologies and a range of projects they initiated or were involved in during the period 1978–1983 aimed to improve Alaska Native peoples' access to public institutions. Through a nexus analysis, they show how the use of computer technology in Alaska changed social practices in educational settings. Originally, this technology was designated to the administrative culture of education, but it proved to be well-suited to creating new learning environments for students who were not comfortable with the panopticon classroom interaction order (the teacher-centered class, where the teacher controls the interaction). Students who had no voice in the classroom setting because the interactional patterns were so different from the cultural codes of their everyday lives used technology as a means for academic discussions on their own terms.

This chapter focuses on some of the ways nexus analysis has been applied to multilingual practices, contexts, and encounters. Nexus analysis is an analytical framework

https://doi.org/10.1515/9783110435351-047

that aims to provide tools for researching language (in a broad sense) as part of social action. The persevering goal of nexus analysis and its several applications is to provide socially relevant results on the ways in which discourse figures in social actions, practices, and processes. In this way, nexus analysis aligns with many other socially and critically oriented qualitative methodological frameworks. While nexus analysis is indeed a form of discourse analysis that combines methodological tools from a wide range of disciplines (including ethnography, interactional sociolinguistics, linguistic anthropology, and social semiotics), it takes *action* rather than *language* as its primary unit of analysis. Within a nexus analytical framework, all actions are understood to be mediated by cultural tools (or mediational means), ranging from language to material objects. Social action, or mediated action, is seen as any action performed by a social actor through the use of mediational means, and the starting point of a nexus analysis is to identify crucial social actions, and then to map the cycles of people, places, discourses, objects, and concepts which circulate through the moment when the social actions take place (Scollon and Scollon 2004). Nexus analysis is, however, not restricted to an analysis of social actions. It is a form of critical language research which not only records actions, but also aims to discuss the involvement of the researchers and results of the analysis with stakeholders, with the broader goal of developing new solutions, practices, and ways of working (Pietikäinen, Compton, and Dlaske 2015). As the next sections will show, social change is both an integral part, and a result, of nexus analysis.

2. Nexus analysis: doing mediated discourse analysis

Nexus analysis can be described as the methodological development of mediated discourse analysis (MDA). Following Scollon and de Saint-Georges (2012: 66) it is "the historical, ethnographic and methodological arm of Mediated Discourse Analysis". According to Norris and Jones (2005: 201) the method that MDA uses "to attempt to answer questions about discourse and social action […] is referred to as *nexus analysis*".

MDA was developed by Ron and Suzie Scollon and some of their colleagues at the turn of the 21th century with a particular focus on linkages between discourse and action, and how these play out in complex social situations (Scollon and de Saint-Georges 2012). A key concept of MDA is social action. All actions are understood to be mediated by cultural tools (or mediational means). The term 'mediational means' was introduced by Wertsch (1991), who underscored that social action and mediational means are intrinsically linked. As such, all human action, even actions we perform when we are alone, is inherently social. Mediational means are semiotic tools ranging from language to material objects. Social action, or mediated action, is seen as any action performed by a social actor through the use of mediational means (Scollon and Scollon 2004). These are the theoretical underpinnings also of nexus analysis, and many researchers who primarily identify MDA as their *theoretical* framework also use nexus analysis as their *methodological* approach. Conversely, when doing nexus analysis, the theoretical underpinnings are shared with MDA. Drawing a clear-cut distinction between MDA and nexus analysis is therefore difficult.

3. Nexus of practice

Rather than taking the speech community as the point of departure, the focus of nexus analysis is social actions and social practices, because it is possible to participate in social practices without belonging to the 'same' community. In research on language contact and multilingualism, there has traditionally been a strong focus on speech communities, but it is important to keep in mind that practices do not require community membership. When a social action is taken repeatedly, it is considered a social practice, which can be understood as an accumulation of social actions. As stated by Scollon (2001:149), a practice "predates the social actor; that is, we mostly learn the practices of our society, but rarely initiate them". In most cases, a practice predates the moment in which the social actor engages in that practice, but as new practices arise, there will be practices that do not precede the social actor. New practices can often arise from individual actions, when these actions are mutually construed as practices within a nexus of practice.

A nexus of practice is the intersection (or nexus) of discourses, social actions and multiple practices that are recognizable to social actors. Scollon and Scollon (2004: 159) define a nexus of practice as "the point at which the historical trajectories of people, places, discourses, ideas, and objects come together to enable some action which in itself alters those historical trajectories in some way as those trajectories emanate from this moment of social action". In contrast to the notion of 'nexus of practice', concepts such as 'community of practice' and 'speech community' presuppose a community. When doing a nexus analysis, the notion of community as the primary analytical unit is not rejected, but rather reserved for when the nexus of practice becomes recognized by social actors, and group membership becomes salient.

Nexus analysis is a form of ethnographic fieldwork because it involves participation and systematic observation, which are two key aspects of ethnography. Both the starting point, and the theoretical and methodological focal point, of study is social action, as opposed to any "a priori social group, class, tribe, or culture" (Scollon and Scollon 2004: 13). As nexus analysis neither excludes, nor presupposes, membership in a social group or community, nexus analysis allows for opening up the scope of multilingualism research, as more fluid social practices also can be studied. Social practices, both discursive and non-discursive, get linked up over time to form a nexus of practice, whereby a nexus of practice is formed one social action at the time (Scollon 2001).

Though nexus analysis does not take *a priori* groups as a starting point, researchers studying multilingual contexts and practices may carry out 'traditional' ethnographic studies with a community as a focal unit (for instance, in studies of language shift, see Lane 2010; Forsman 2015), study groups or communities of practice (e.g. investigating the challenges faced by Latino educators in a teacher education program during an educational reform, see Austin et al. 2010), or conduct studies of multilingual practices not grounded in membership of a particular group or community (notably in linguistic landscape studies; Hult 2014). A nexus of practice can also be quite fleeting, as suggested by Pennycook and Otsuji (2017), for instance, during encounters in shops, markets, waiting rooms at a medical center, and so on.

4. Interactional order, historical bodies, and circulating discourses: three main elements of nexus analysis

Scollon and Scollon (2004: 8) state that the aim of nexus analysis is to unify two strands of discourse analysis by taking into consideration both the "micro-analysis of unfolding moments of social actions or a much broader socio-political-cultural analysis of the interrelationships among social groups and power interests in society". Social issues are seen as being grounded in the micro-level actions of individuals (i.e. in a moment of language use), and larger structures and processes of social organization frame the moment where language use takes place. Social actions occur at a concrete moment in space and time, and are closely linked to the histories of the social actors, social interactions, and discourses.

Nexus analysis takes a social action performed by a social actor as the starting point of analysis. As illustrated in Figure 47.1, a social action occurs at the intersection of three key elements: the historical bodies of the participants, the interaction order (mutually negotiated and coproduced by the participants), and discourses in place (which enable the action and/or are used by participants as mediational means; Scollon and Scollon 2004: 153).

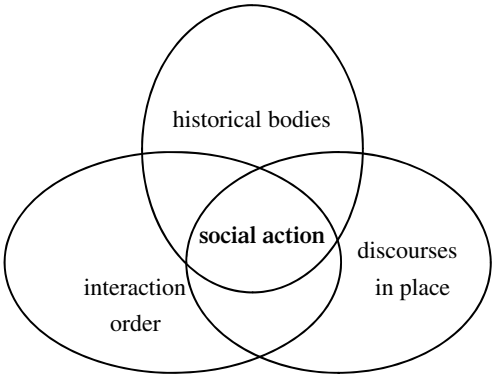


Fig. 47.1: Main elements of nexus analysis (Based on Scollon and Scollon 2004: 154)

4.1. Interaction order

Goffman (1983) noted that individuals behave differently when they are with others adhering to conventionalized norms, and introduced the term interaction order to account for how social actors form relationships in social interactions. He was particularly concerned with "social situations, that is, environments in which two or more individuals are physically in one another's response presence" (Goffman 1983: 2), and named this face-to-face domain the interaction order. Scollon and Scollon (2004:13) point out how

social actions partially depend on whether we are alone or with others, and what setting we are in: "[o]ur roles and role expectations depend on the relationship and interactions between social actors, in short − the social arrangements by which we form relationships in social interactions".

4.2. The historical body

The historical body can be seen as the life histories of the social actors, or the embodied biographies and life experiences of speakers. Our historical body is the lifetime accumulations of our actions, memories, and experiences, defined by Scollon and Scollon (2004: 13) as follows:

> Different people play the same role differently depending on their history of personal experience inscribed in what the philosopher Nishida calls the historical body. A lifetime of personal habits comes to feel so natural that one's body carries out action seemingly without being told. Bourdieu referred to this phenomenon as habitus but we prefer historical body because it situates bodily memories more precisely in the individual body.

Social practices may become internalized and integrated in our historical bodies to such an extent that we are no longer aware of them. The term historical body encompasses the notions of embodiment and time, concepts that have been brought to the forefront of recent sociolinguistic and discourse analytical research (Blommaert 2010; Busch 2012): the object of study − social action and practice − is situated in real time, but produced by historical processes (Lane 2014).

4.3. Discourses in place

Social actions are carried out at specific, concrete moments and places in the material and physical world, and discourses on different scales circulate through these nexuses. Some discourses, such as national language policy processes, circulate on long time scales and may seem distant to the social actions we are observing. When materialized in textbooks, for instance, such discourses become more directly relevant to social actions and actors, as they may both constrain and facilitate action. Through engagement in the nexus of practice, the researcher aims to map the discourses that circulate through the moment in which the action occurs, and to identify which discourses are foregrounded versus which seem irrelevant.

In the work of Hult (2010, 2017), nexus analysis is considered a scalar approach to language policy and to ethnography of educational linguistics. Hult (2010: 14) defines scale as "a sociolinguistic construct that refers to the fluid and dynamic nature of relationships among discourse processes across dimensions of social organization. Rather than characterizing social phenomena in terms of micro-macro layers, the scale construct foregrounds a view of discourse processes as situated in time and space, or 'TimeSpace' (Blommaert 2007: 5)". This is also brought to the fore by researchers such as Pietikäinen (2014: 482), who draws on Bahktin's concept of chronotope in her analysis of the multilingual landscape in a Sámi village in Northern Finland, stating that signs "function as

primary means by which time is materialized in space: temporality and spatiality are material conditions. They can be seen as patterns of flow, exchange and interaction". All social action and discourses are historically situated, with trajectories originating from the past and pointing forward to possible futures. As pointed out by Scollon and Scollon (2004: 164), it is difficult to see 'invisible' discourses during an event in the present, because these discourses tend to be visible only by mapping discourse cycles backward (or forward) from the event one is analyzing.

5. Activities of nexus analysis: engaging, navigating, and changing the nexus of practice

A characteristic of nexus analysis in multilingualism research is an engagement with social issues. Researchers doing nexus analysis in multilingual contexts have been concerned with issues such as school and home relations (Bergroth and Palviainen 2016), language and pedagogy (Jocuns 2007; Levine 2015; Pietikäinen and Pitkänen-Huhta 2013), and language shift (Forsman 2015; Lane 2010). The initial stage of a nexus analysis involves establishing the social issue one wants to study, finding the crucial social actors, and identifying the crucial social actions where that social issue is produced, ratified, or contested. Having identified the crucial social actions, the next steps are the three key activities of a nexus analysis: engaging, navigating, and changing the nexus of practice.

5.1. Engaging the nexus of practice

Engaging the nexus of practice involves identifying a social issue and key social actions, finding the primary social actors, observing the interactional order, and identifying significant cycles of discourse. In order to achieve this, the researcher seeks to become involved with the participants of the nexus of practice in order to identify the social actions, discourses, and mediational means that are relevant to the participants. Scollon and Scollon (2004) refer to this as establishing a 'zone of identification', which is the initial stage of any ethnographic fieldwork (Agar 1996). The first task of a nexus analysis is thus to determine what kinds of data should be gathered, and how to gather it. A nexus analyst seeks to identify relevant and important social actions. Social actions can occur on different scales. Norris (2004) suggests that both higher- and lower-level actions should be part of an analysis focusing on social action. A higher-level action may be defined as "an action that is bracketed by an opening/closing and made up of a multiplicity of chained lower-level actions. Often, we find several higher-level actions embedded in another, and/or overarching higher-level action" (Norris and Maier 2014: 388, with reference to Pirini, who provided the definition). Norris (2004: 11) defines a lower-level action "as the smallest interactional meaning unit". Interviewing may then be seen as a social practice, and a specific interview as a higher-level social action which in turn consists of chains of lower-level social actions, such as posing and answering a question, turn-taking, pauses, gaze, etc. These social actions may also overlap, so that, for example, gaze and posing a question often occur simultaneously.

Gathering and analysis of data are not discrete processes: researchers analyze (parts of) the data during the process of gathering data, which in turn makes them see what other kinds of data might be needed.

5.2. Navigating the nexus of practice

The main activity of a nexus analysis is to map the cycles of people, places, discourses, objects, and concepts circulating through the moment in which a social action takes place (Scollon and Scollon 2004: 159). Scollon and Scollon (2004: 160) describe the relationship between the social action and the semiotic cycles as follows:

> If we think of an action as a moment in time and space in which the historical bodies and the interaction order of people and the discourses in place intersect, then each of these can be thought of as having a history that leads into the moment and a future that leads away from it in arcs of semiotic cycles of change and transformation.

This activity involves the gathering of data and analysis of selected data. Often, these activities are not separate processes, but rather are carried out at the same time. Ideally, the researcher, informed by interaction with participants, will narrow down the scope of the analysis, and select practices, discourses, social actions, and/or sequences of actions to be analyzed in more detail.

Doing a nexus analysis implies doing research with people. Scollon and Scollon (2004: 153) underscore that there is no study from afar in nexus analysis (or in any ethnographically oriented research), and the aim is for the researcher to be recognized by other participants as a participant in the nexus of practice with which they have engaged. This does not necessarily require that the researcher is recognized as an in-group member, but rather that the researcher develops a role for themselves in order to participate in the nexus of practice. This is especially true in the case of nexus analysis with a participant-observation focus, but the role of the researcher will be different in a nexus analysis of, for example, policy texts or linguistic landscape studies.

The presence of a researcher has been frequently discussed, and to some extent cast, as a methodological problem for sociolinguistics when interrogating how we can observe the unobserved. Within nexus analysis, as in most critical approaches, the presence of the fieldworker is not seen as a problem, but rather as an asset, when it comes to identifying relevant social actions, practices, discourses, and mediational means (Lane 2014). Some nexus studies are more grounded in participant-observation, whereas others are more discourse-focused; many nexus analysis studies combine both aspects. Whether a nexus analysis is a 'full-fledged' ethnographic study or not, such a study is always informed by ethnographic awareness and reflexivity.

Researchers working within a nexus analytical framework draw on methods from a wide range of fields, such as discourse analysis, linguistic anthropology, narrative analysis, literacy studies, education studies, and social semiotics. As mediational means encompasses both semiotic and material tools, many nexus analysis studies combine different types of data and focus on different modes. Throughout these activities, the researcher, guided by interaction with participants, will narrow down the scope of the analysis and select a practice, a discourse, a social action, or a sequence of actions.

5.3. Changing the nexus of practice

When our data come through engagement in the nexus of practice (as found, for example, in ethnographic fieldwork), the very presence of the researcher to some extent becomes part of the object of study, as it is impossible to abstract away from the researcher's presence. Their presence, actions, and historical body will influence the interaction order and discourses in place. Ethnography and sociolinguists have struggled with this paradox. Labov (1972: 209) formulated the Observer's Paradox, or how to observe the unobserved: "[t]he aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain these data by systematic observation". When doing a nexus analysis, the researchers' presence is recognized as an integral part of the analysis, because they must become involved with the people they are working with in order to identify the social actions, discourses, and mediational means that are relevant to the participants themselves. Through this process of establishing a zone of identification, the researcher becomes a part of the nexus of practice (at least to some extent), and thereby also one of its influencing factors.

As previously stated, the starting point of a nexus analysis is a social issue, and therefore the engagement of the researcher will influence and change practices, and the nexus of practice itself. This is a two-way process, as this engagement will also change the researcher, which may in turn have an impact on the choice of research questions, methods, and theoretical insights; being part of the analysis also influences the historical body of the researcher. Hence, the change brought about by a nexus analysis is not necessarily determined in advance, but develops as an outcome of the researcher's engagement with a nexus of practice (Lane 2014). Nexus analysis takes the analysis, interpretation, and explanation of social issues as its central concern, and as a result, researchers applying nexus analysis to contexts and encounters often address social issues in their research.

5.4. Nexus analysis: an example

To illustrate the three central activities of a nexus analysis, an exemplary study is presented (Pan 2014). While this is an example of a study that encompasses the three activities of a nexus analysis, it is important to bear in mind that all nexus analysis-based studies will be different. Norris and Jones (2005: 201) point out that nexus analysis is not a strict methodology, but rather a set of heuristic tools that guide researchers: each activity is "a kind of a reflective process, involving a multitude of contingencies and constant interactive decision making on the part of the researcher". As such, the study presented here is an example of what a nexus analysis *might* look like, but not necessarily what a nexus analysis *should* look like.

Pan (2014) presents a nexus analysis study that her research team carried out based on the 2010 United States census, in which the US Census Bureau conducted a nationwide census consisting of two stages: a written census form, and in-person interviews with those who did not return the census forms mailed to them. A multilingual research team observed census interviews conducted with respondents who were speakers of seven non-English languages and/or multilingual speakers. As is the case for much research

on multilingualism, a nexus analysis is influenced, and sometimes even motivated, by an engagement with a social issue or an "issue that has some social implications" (Pan 2014: 56). The research team was concerned that, because non-English speaking populations are likely to be overlooked in a census counting, this might affect the political representation of minority groups (i.e. the social issue of interest). When *engaging the nexus of practice*, the researchers identified the higher-level action (in Norris' terms) to be studied as the act of enumerating non-English-speaking populations through face-to-face interviews. They identified crucial social actors as the non-English speaking respondents and the census takers, and outlined how the researchers gained a place in the nexus of practice by spending time with the local census offices to become familiar with census procedures and requirements, and by accompanying census takers on interviews for two weeks prior to gathering data.

Pan (2014) describes *engaging* the nexus of practice as the ethnographic component of the research process, and *navigating* the nexus of practice as largely a discourse analysis of central and crucial cycles of discourse. Yet as for ethnographic and discourse analytical studies in general, there is no clear-cut delineation between engaging with research participants, and the process of gathering and analyzing data. When navigating the nexus of practice, Pan and her colleagues sought to link the census event (and the broader discourse of the US census) with immigrants' experiences with census-taking in their country of origin and in the US (and their perceptions of census-taking as a social practice). In order to investigate this, they took detailed notes during the interviews and analyzed recordings, and included data about the clothing worn by participants, texts and materials used during interviews, place and location of interviews, and signs, images, and graphics. After the census interview, participants were asked about their experience to see if their perception of the US census matched the social practice of census-taking in the US. The analysis showed that many immigrants had little understanding of the purpose and procedures of the US census, and that many were reluctant to participate in the census (see Pan 2014).

Change is an integral part of nexus analysis, although the degree of a researcher's activist position varies. In the exemplary project presented here (Pan 2014), the goal of positive change, or *changing* the nexus of practice, was an integral part of the study. The research team brought their findings and observations back to the community they observed, and at the conclusion of the project, the research team made extensive recommendations about how to improve the census process, and their findings were discussed at various levels of the census operation. Not all nexus analysis projects take such an explicit stance when it comes to promoting positive change, but for those who do, engaging the nexus of practice is essential, as change is something that should ideally arise from the nexus of which the researcher is but one participant; in other words, the researcher and the participants make change come about *through their mutual engagement* in the nexus of practice.

6. Language contact: a multilingual and multimodal turn

At the core of contemporary research on multilingualism, there is an understanding that our social world is complex, dynamic, and diverse. Our field has undergone a methodo-

logical and epistemological shift in recent decades. New strands of research on multilingualism have emerged, "which have incorporated critical and post-structuralist perspectives from social theory and embraced new epistemologies and research methods. There has also been a shift of focus to empirical work, which is interpretive, ethnographic and multimodal in nature" (Martin-Jones, Blackledge, and Creese 2012: 1). As such, there is a need for methods that can address these complexities, and there is a growing awareness that we need theories and methods that can allow us to focus not only on language in a narrow 'linguistic' sense, but also on situating social actors at the heart of the analysis (Kramsch 2009). Block (2014: 56) calls for a wider scope of research on multilingualism through "active engagement with embodiment and multimodality as a broadened semiotically based way of looking at what people do when they interact".

Nexus analysis is a methodology suitable for analyzing complex social phenomena and processes precisely because one of the primary concerns of nexus analysis is the complex relationship between discourse and action: it places the experiences of individual social actors at the heart of the analysis. This might be one of the reasons why researchers who are interested in modes other than language only in a 'linguistic' sense combine nexus analysis with multimodal approaches. Nexus analysis has been fruitfully applied in linguistic landscape studies (see for instance Hult 2014; Lou 2016; Pietikäinen et al. 2011). Drawing on Norris and Jones' (2005) concept of frozen actions, Pietikäinen et al. (2011) analyzed signs in the linguistic landscapes of multilingual areas in the Far North as the material results of actions carried out in the past, while Lane (2017) explored texts as the material results of minority language standardization processes. Researchers doing nexus analysis in the context of education also frequently draw on multimodal methods (see e.g. Pietikäinen and Pitkänen-Huhta's [2013] analysis of classroom literacy practices in indigenous multilingual settings, wherein they investigate the making of a picture book as a multimodal and visual activity). This concern with materiality is also evident in Izadi (2017: 465), who draws on the concept of "semiotic aggregates", or the intersection of multiple semiotic systems (including objects), to investigate local language practices in two shops in a multilingual area in Sydney, arguing that "relying solely on spoken language, without paying much attention to the importance of the social processes, the historical and cultural context that underpins such verbal exchanges, would lead to an incomplete picture of the instances and the structure of such interactions". Pennycook and Otsuji (2017) focus on assembling artefacts (objects that mediate across geographies, environments, culinary traditions, and histories) in two Bangladeshi-run stores in Sydney and Tokyo, showing that such objects must be taken very seriously as part of the action. Central to the study is the way(s) in which objects come together as part of semiotic assemblages of material and semiotic resources, as customers, goods, and languages assemble (and disassemble) at particular moments. The authors argue that this has implications for how we understand the role of objects in social semiotics as part of a critical sociolinguistics of diversity: any critical sociolinguistics of diversity must encompass a diversity of assembled human and nonhuman artefacts.

7. Conclusion: what type of research is nexus analysis suited for?

Ron and Suzie Scollon's book *Nexus analysis: Discourse and the Emerging Internet* was published in 2004; at that time, a type of discourse analysis that did not privilege lan-

guage, but which saw the utterance as but one example of a mediated action, was not uncontroversial because language was seen as the privileged unit of analysis for sociolinguistics and discourse analysis (Norris 2010: 32). Recent research in these fields has been shaped by poststructuralist theories, emphasizing pluralism, hybridity, and human agents as co-constructors of language and culture, and a shift towards seeing language as a social object and investigating globalization and transnational flows (Lane 2012). The world we live in and research on multilingualism have both become increasingly multimodal, technologically saturated, and complex, and it has been suggested that we experience an era of superdiversity (Vertovec 2007). Though the novelty of this notion has been contested (see for instance Pavlenko 2019), researchers engaging with multilingualism argue that we need theoretical approaches that take complexities into account. There has also been a growing awareness that language constitutes a part of a larger social semiotic system, as underscored by Block (2014: 56), who argues for "an active engagement with embodiment and multimodality as a broadened semiotically based way of looking at what people do when they interact". The notion of embodiment is also brought to the fore by Kramsch's (2009) study of the embodied and emotional dimensions of language learning, Busch's (2012) research on language repertoires, and Pujolar and Gonzàlez' (2013) focus on specific moments (or life events) of individuals that led them to position themselves as speakers of a language which had not previously been part of their linguistic repertoire. Embodiment is also evident in nexus analysis' concept of historical bodies, which underscores the lived experiences and memories of social actors.

Nexus analysis emphasizes that discourses and individual social actors have a history, and that these therefore cannot be analyzed without reference to the past. The time perspective is thus crucial: people, objects, mediational means, and discourses are seen as having a history and projecting a future (de Saint Georges 2005). This framework is therefore ideally suited to analyses of social processes that work on longer time scales, e.g. nexus analyses of language shift (Forsman 2015; Lane 2010). As pointed out by Scollon and Scollon (2004), nexus analysis is a form of ethnographic discourse, and can also be seen as a historical discourse analysis, as it incorporates the notion of time and discourse itineraries. Pietikäinen (2015) takes a rhizomatic approach to nexus analysis by highlighting the interconnectedness of discourses: seeing processes and events in terms of flow and (dis)connections, and suggesting that one of the aims of nexus analysis is to trace the lines of trajectories which connect different discourses.

Nexus analysis does not focus primarily on language in and of itself, but rather on how language and other semiotic and/or material tools are used to mediate action, albeit without assigning a privileged position to language as compared to other mediational means. Nexus analysis is therefore primarily suited for qualitative research on multilingualism, and in particular, for investigating social practices and incorporating multimodal approaches.

Researchers who do nexus analysis are often motivated by a concern with social issues, described by Scollon and Scollon (2007: 608) as follows: "[n]exus analysis is our response to Hymes' call three decades ago for each of us to reinvent anthropology as 'a personal general anthropology, whose function is the advancement of knowledge and the welfare of mankind'". This statement situates nexus analysis firmly in the critical tradition of the social sciences and humanities, which combines academic research with

a social (and often political) engagement (Wodak 2006). Nexus analysis, as all ethnographic research, will lead to change, both on the immediate and longer time scales. A researcher's engagement in the nexus of practice will bring about change, and this is a multifaceted process, as such engagement will also become a part of the historical body of the researcher. In their introduction to a special issue in memoriam of Ron Scollon, Lou and Stockburger (2010: 4), described Ron Scollon as "a scholar who spent his career and life engaged with social issues in the real world and, by his very own example, made us feel hopeful that research on language and discourse can indeed contribute to positive social change". This concern with social issues and language in the real world is at the very heart of nexus analysis.

Acknowledgements

This work was partly supported by the Research Council of Norway through its Centres of Excellence funding scheme, namely the projects *Centre for Multilingualism in Society across the Lifespan* (#223265) and *Standardising Minority Languages* (#213831). I would also like to thank Sari Pietikäinen, who read an early draft of this chapter and gave invaluable feedback and suggestions.

8. References

Agar, Michael H.
1996 *The professional stranger: An informal introduction to ethnography*, 2nd edn. San Diego: Academic Press.

Austin, Theres, Jerri Willett, Margaret Gebhard & Agustin Laó Montes
2010 Challenges for Latino educators crossing symbolic, cultural, and linguistic boundaries: Coming to voice in teacher preparation with competing voices. *Journal of Latinos and Education* 9(4). 262–283.

Bergroth, Mari & Åsa Palviainen
2016 The early childhood education and care partnership for bilingualism in minority language schooling: Collaboration between bilingual families and pedagogical practitioners. *International Journal of Bilingual Education and Bilingualism* 19(6). 649–667.

Block, David
2014 Moving beyond "lingualism": Multilingual embodiment and multimodality in SLA. In Stephen May (ed.), *The multilingual turn: Implications for SLA, TESOL, and bilingual education*, 32–77. New York: Routledge.

Blommaert, Jan
2007 Sociolinguistic scales. *Intercultural Pragmatics* 4(1). 1–19.

Blommaert, Jan
2010 Historical bodies and historical space. *Working Papers in Urban Language & Literacies* 57. https://www.kcl.ac.uk/sspp/departments/education/research/research-centres/ldc/publications/workingpapers/the-papers/57.pdf (accessed on 3 January 2019).

Busch, Brigitta
2012 The linguistic repertoire revisited. *Applied Linguistics* 33(5). 503–523.

de Saint Georges, Ingrid
2005 From anticipation to performance: Sites of engagement as performance. In Sigrid Norris & Rodney H. Jones (eds.), *Discourse in action: Introducing mediated discourse analysis*, 155–165. London: Routledge.

Forsman, Ludvig
2015 Language shift in Gammalsvenskby: A nexus analysis of the shift to Ukrainian in a traditionally Swedish-speaking community. *Journal of Sociolinguistics* 19(1). 37–61.
Goffman, Ervin
1983 The interaction order: American Sociological Association, 1982 Presidential address. *American Sociological Review* 48(1). 1–17.
Hult, Francis
2010 Analysis of language policy discourses across the scales of space and time. *International Journal of the Sociology of Language* 202. 7–24.
Hult, Francis
2014 Drive-thru linguistic landscaping: Constructing a linguistically dominant place in a bilingual space. *International Journal of Bilingualism* 18(5). 507–523.
Hult, Francis
2017 Nexus analysis as scalar ethnography for educational linguistics. In Marilyn Martin-Jones & Deirdre Martin (eds.), *Researching multilingualism: Critical and ethnographic perspectives*, 89–104. New York: Routledge.
Izadi, Dariush
2017 Semiotic resources and mediational tools in Merrylands, Sydney, Australia: The case of Persian and Afghan shops. *Social Semiotics* 27(4). 495–512.
Jocuns, Andrew
2007 Semiotics and classroom interaction: Mediated discourse, distributed cognition, and the multimodal semiotics of Maguru Panggul pedagogy in two Balinese Gamelan classrooms in the United States. *Semiotica: Journal of the International Association for Semiotic Studies / Revue de l'Association Internationale de Sémiotique* 164(1–4). 123–151.
Kramsch, Claire
2009 *The multilingual subject.* Oxford: Oxford University Press.
Labov, William
1972 *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
Lane, Pia
2010 "We did what we thought was best for our children": A nexus analysis of language shift. *International Journal of the Sociology of Language* 202. 63–78.
Lane, Pia
2012 Multimodality and culture. In Carol A. Chapelle (ed.), *Blackwell encyclopedia of applied linguistics*. Oxford: Blackwell. https://doi.org/10.1002/9781405198431.wbeal0821. (accessed on 11 September 2018).
Lane, Pia
2014 Nexus analysis. In Jan-Ola Östman & Jef Verschueren (eds.), *Handbook of pragmatics*, vol. 18. Amsterdam: John Benjamins. https://doi.org/10.1075/hop.18.nex1. (accessed on 11 September 2018).
Lane, Pia
2017 Language standardisation as frozen mediated actions: The materiality of language standardisation. In Pia Lane, James Costa & Haley De Korne (eds.), *Standardizing minority languages: Competing ideologies of authority and authenticity in the global periphery*, 101–117. New York: Routledge.
Levine, Glenn
2015 A nexus analysis of code choice during study abroad and implications for language pedagogy. In Jasone Cenoz & Durk Gorter (eds.), *Multilingual education: Between language learning and translanguaging*, 84–113. Cambridge: Cambridge University Press.
Lou, Jackie Jia
2016 *The linguistic landscape of Chinatown: A sociolinguistic ethnography.* Bristol, UK: Multilingual Matters.

Lou, Jackie Jia & Inge Stockburger (eds.)
2010 *A special issue in honor of Ron Scollon. eVox Georgetown Working Papers in Language, Discourse, and Society* 4(1).
Martin-Jones, Marilyn, Adrian Blackledge & Angela Creese (eds.)
2012 *The Routledge handbook of multilingualism*. London: Routledge.
Norris, Sigrid
2004 *Analyzing multimodal interaction: A methodological framework*. New York: Routledge.
Norris, Sigrid
2010 Mediated discourse and social interaction: A reflection. *eVox Georgetown Working Papers in Language, Discourse, and Society* 4(1). 32–33.
Norris, Sigrid & Carmen Daniela Maier (eds.)
2014 *Interactions, images and texts: A reader in multimodality.* Berlin: De Gruyter Mouton.
Norris, Sigrid & Rodney H. Jones (eds.)
2005 *Discourse in action: Introducing mediated discourse analysi*s. London: Routledge.
Pan, Yuling
2014 Nexus analysis. In Sigrid Norris & Carmen Daniela Maier (eds.), *Interactions, images and texts: A reader in multimodality*, 53–62. Berlin: Mouton De Gruyter.
Pavlenko, Aneta
2019 Superdiversity and why it isn't: Reflections on terminological innovation and academic branding. In Stephan Breidbach, Lutz Küster & Barbara Schmenk (eds.), *Sloganizations in language education discourse*, 142–168. Bristol, UK: Multilingual Matters.
Pennycook, Alastair & Emi Otsuji
2017 Fish, phone cards and semiotic assemblages in two Bangladeshi shops in Sydney and Tokyo. *Social Semiotics* 27(4). 434–450.
Pietikäinen, Sari
2014 Spatial interaction in Sámiland: Regulative and transitory chronotopes in the dynamic multilingual landscape of an indigenous Sámi village. *International Journal of Bilingualism* 18(5). 478–490.
Pietikäinen, Sari
2015 Multilingual dynamics in Sámiland: Rhizomatic discourses on changing language. *International Journal of Bilingualism* 19(2). 206–225.
Pietikäinen, Sari & Anne Pitkänen-Huhta
2013 Multimodal literacy practices in the indigenous Sámi classroom: Children navigating in a complex multilingual setting. *Journal of Language, Identity & Education* 12(4). 230–247.
Pietikäinen, Sari, Pia Lane, Hanni Salo & Sirkka Laihiala-Kankainen
2011 Frozen actions in the Arctic linguistic landscape: A nexus analysis of language processes in visual space. *International Journal of Multilingualism* 27(1). 1–23.
Pietikäinen, Sari, Sarah E. Compton & Kati Dlaske
2015 Putting resources into practice: A nexus analysis of knowledge mobilisation activities in language research and multilingual communities. *Current Issues in Language Planning* 16(3). 187–200.
Pujolar, Joan & Isaac Gonzàlez
2013 Linguistic 'mudes' and the de-ethnicization of language choice in Catalonia. *International Journal of Bilingual Education and Bilingualism* 16(2). 138–152.
Scollon, Ron
2001 Action and text. In Ruth Wodak & Michael Meyer (eds.), *Methods of critical discourse analysis*, 139–183. London: Sage.
Scollon, Ron & Suzie Wong Scollon
2004 *Nexus analysis: Discourse and the emerging internet*. London/New York: Routledge.

Scollon, Ron & Suzie Wong Scollon
2007 Nexus analysis: Refocusing ethnography on action. *Journal of Sociolinguistics* 11(5). 608–625.
Scollon, Suzie & Ingrid de Saint-Georges
2012 Mediated discourse analysis. In James Paul Gee & Michael Handford (eds.), *The Routledge handbook of discourse analysis*, 66–78. London: Routledge.
Vertovec, Steven
2007 Super-diversity and its implications. *Ethnic and Racial Studies* 30(6). 1024–1054.
Wertsch, James
1991 *Voices of the mind: A sociocultural approach to mediated action.* Cambridge, Mass.: Harvard University Press.
Wodak, Ruth
2006 Critical linguistics and critical discourse analysis. In Jan-Ola Östman & Jef Verschueren (eds.), *Handbook of pragmatics*, vol. 10. Amsterdam: John Benjamins. https://doi.org/10.1075/hop.10.cri1. (accessed on 31 December 2018).

Pia Lane, Oslo (Norway)

48. Ethnography in research on language contact

1. Introduction
2. Ethnography in sociolinguistic research: Epistemic and disciplinary foundations
3. Practical considerations in ethnographic research
4. Illustrative ethnographies of language contact phenomena
5. Challenges and opportunities
6. References

1. Introduction

As the field of contact linguistics has grown, so too have the methodological approaches used to understand language contact phenomena. Language contact is a "multidimensional, multidisciplinary field", writes Clyne (2003: 1), and languages in contact "are, after all, the result of people in contact and of communities of people of different language backgrounds in contact". It is the everyday, 'peopled' aspects of language contact that ethnography is ideally suited to study.

Ethnography is both a *process* of conducting research framed within a particular ontological, epistemic, and methodological tradition, and a *product* (or account) derived from that process. Rooted in the discipline of anthropology, ethnography is a 'way of seeing' through the lens of human culture, and a 'way of looking' based on situated, long-term, in-depth fieldwork (Wolcott 2008). As a way of *seeing*, ethnography is a cultural interpretation of the meanings people make of everyday social practice (Anderson-Levitt 2006). As a way of *looking*, ethnographic methods include participant and nonparticipant observation; informal, semi-structured, and structured interviews; and the collection and

https://doi.org/10.1515/9783110435351-048

analysis of documents and other cultural artifacts. These data-gathering activities are often supplemented with surveys, photographs, online artefactual sources, and audio and/or video recordings.

Ethnography can also be seen as a 'way of being' a researcher that is humanizing, reflexive, and counter-hegemonic (McCarty 2015b: 81). For instance, in contact linguistics, the ethnographer might ask how and why certain language varieties (such as dominant 'standard' English in the US) are privileged in the workplace, education, government, media, and technology, while other languages and varieties (and their speakers) are stigmatized. As Blommaert and Jie (2010: 11) point out, such "mapping of resources" onto language statuses and functions is not a property of language, but rather of society; ethnography thus becomes a "critical enterprise" of deconstructing the power relations that (re)produce linguistic hierarchies. For many ethnographers, this entails a commitment to use ethnographic knowledge to transform linguistic inequalities. "The purpose of research is not just […] describing the world", says Lin (2015: 26), "but also changing the world" (see also Hymes 1974, 1980, 1996).

In this chapter, we explore the use of ethnography in the study of language contact phenomena, and the contributions, challenges, and value it brings to the field. Our discussion is guided by the early work of scholars such as Joshua Fishman, Einar Haugen, and Uriel Weinreich, who positioned language contact phenomena as fundamentally sociocultural in nature. We are also guided by recent work that embraces an expansive view of contact linguistics, and takes a critical, humanizing, and decolonizing stance toward qualitative inquiry (Paris and Winn 2014; Smith 2012). While pidgins and creoles are "the cornerstone of the study of languages in contact" (Lim and Ansaldo 2016: x), language contact also entails a host of issues surrounding bi/multilingualism and education, code-mixing and translanguaging, communication in the digital world, diverse Englishes, globalization and superdiversity, linguistic landscapes, and language shift and reclamation. (For key sources, see Androutsopoulos 2006; Arnaut et al. 2016; García and Li Wei 2014; Hinton, Huss, and Roche 2018; Martin-Jones, Blackledge, and Creese 2012; Onysko 2016; Shohamy and Gorter 2009). Although space prohibits a full exploration of the range of this work, in Section 4 of this chapter, we highlight key ethnographic studies.

We begin by sketching the evolution and use of ethnography in sociolinguistic research, and how ethnography is realized in contemporary practice. We then discuss ethnography's challenges and its value for future research on language contact.

2. Ethnography in sociolinguistic research: epistemic and disciplinary foundations

The term ethnography derives from the Greek *ethnoi*, 'the nations, the others', and *graphein*, 'to write' (Erickson 2011: 45). This suggests the colonialist origins of ethnographic research, as "anthropologists began to use the term *ethnography* for descriptive accounts of the lifeways of […] people who lived in colonial situations around the world" (Erickson 2011: 44).

Contemporary ethnography is often traced to the early 20th century influence of Franz Boas in North America and Bronisław Malinowski in the UK. From the Boasian school

of American anthropology came an emphasis on cross-cultural comparison and cultural relativism (i.e. the notion that each cultural group must be understood in terms of its own norms and practices), as well as inductive reasoning based on "the empirical gathering of descriptive and historical data" (Baker 1998: 119). Boas and his students applied the principle of relativity to linguistic systems as well, insisting on studying them in terms of their internal properties. From Malinowski came a parallel emphasis on long-term fieldwork and understanding parts of a cultural system (e.g. communicative practices) within the social system as a whole. Both Boas and Malinowski advocated knowledge of the local language to "accurately [represent] nuances of local meaning" (Erickson 2011: 46). "[R]eporting that included the meaning perspectives of those whose daily actions were being described", Erickson (2011: 46) writes, "became a hallmark of ethnography in anthropology".

While early anthropologists interpreted culture as a unitary, bounded system, often reduced to racialized groups and/or traits, more recent work emphasizes cultural processes, discourses, ideologies, and practices. González (2001: xxii) points out that the latter view expands "our vision […] to include issues of power and legitimation, as well as the language practices that constitute these issues". From an ethnographic perspective, then, language, like culture, is an open, dynamic system inextricably linked to human social life and power relations; "language *within* social process is the focus" (González 2001: xxii).

In the 1960s, Gumperz and Hymes brought ethnography into the then-emerging field of sociolinguistics, introducing the ethnography of communication as "studies ethnographic in basis, and of communication in the […] kind of patterned complexity with which [anthropologists] deal" (Hymes 1964: 2). It follows that ethnographic studies of language contact phenomena are deeply contextualized and conducted *in situ* over extended periods of time, centering the meanings participants themselves make of discursive processes and communicative events. The latter is often referred to as the 'emic' perspective, in reference to an analogy proposed by Pike (1967) that contrasted phonemics (the *tacit knowledge* of a sound system possessed by native speakers) with phonetics (the *study* of those sound systems). The terms *emic* and *etic* are commonly understood to refer to 'insider' and 'outsider' knowledge, respectively.

Hymes was particularly concerned with languages in contact − or more accurately, *speakers* of different languages and language varieties in contact − within the realm of education. With the goal of illuminating diverse "ways of speaking" (Hymes 1980: 20), the ethnography of communication reflected a "socially realistic linguistics" in which education was "a prime arena for sociolinguistic research" (Hornberger 2003b: 245–246). Out of this early sociolinguistic work came classic accounts, such as Philips' (1983) microethnography of the divergent communicative interactions experienced by Native American children in the classroom and their community, and Heath's (1983) study of culturally patterned differences in child language socialization in White and Black working class communities in the southeastern US. These and other sociolinguistically oriented studies demonstrated the culturally specific ways in which talk is organized, and how "dominant-subordinant relationships are formed in face-to-face interaction" (Philips 1983: xvi).

A growing ethnographic literature in applied and educational linguistics (Bigelow and Ennser-Kananen 2015; Spolsky and Hult 2008) and the ethnography of language policy (Hornberger et al. 2018; Johnson 2013; McCarty 2011) builds on this early work

in the ethnography of communication. A central stream of this research investigates the processes of language shift and reclamation, exemplified in 'ground up' studies of Indigenous-language reclamation (Leonard and De Korne 2017), research on the 'new speaker' movement in the EU and the UK (O'Rourke and Walsh 2018), and numerous book-length monographs (for a state-of-the-art review, see Hornberger et al. 2018). Summarizing the thrust of this scholarship, Tollefson (2013: 27) notes that a "renewed emphasis" on language practices and ideologies at the local level has led researchers to turn "increasingly to ethnography as a major research method".

3. Practical considerations in ethnographic research

Like all research approaches, ethnography can address some questions, but not others; specifically, ethnography fleshes out the details of language ideologies, beliefs, and practices in (re)organizing human social life. Geertz (1973) refers to this as the 'complex specificness' and 'circumstantiality' of discourses, interactions, activity, and behaviors. Blommaert (2013) characterizes ethnographic accounts of sociolinguistic systems as 'chronicles of complexity'.

Reflecting that complexity, ethnographic research questions are open-ended, and as such, designed to describe and investigate multidimensional sociolinguistic phenomena, not to reduce them to 'yes-no' or cause-and-effect responses. For example, in an ethnographic study of language policy and education reform in Bolivia, researchers asked how faculty and students in a graduate program in bilingual intercultural education constructed and negotiated "what it means to be *indigenous*" (Hornberger and Johnson 2007: 521). In a study of Indigenous youth language practices in the southwestern US, Romero-Little et al. (2007: 609) asked what role the Indigenous language plays in Native American youths' personal, family, community, and school lives. Ethnographic research questions such as these are subject to change, and emerge from careful decision-making informed by disciplinary knowledge, ethical imperatives, and face-to-face interactions in the field.

Ethnographic research often has a praxis dimension: researchers are concerned with transforming ethnographic insights into knowledge of practical use by and for the benefit of the people and communities with whom they work. For instance, in a study of Latino/a children learning English in Arizona (an officially 'English-only' state in the US), Combs, González, and Moll (2011) used ethnographic methods to examine how a bilingual teacher constructed an unofficial, *de facto* language policy wherein students' Spanish-language abilities were employed as resources for learning. These ethnographic understandings suggested the ways in which teaching and learning in bi/multilingual contexts can "challenge, and perhaps even transform, the status quo" (Combs, González, and Moll 2011: 200).

In focusing ethnographic research questions, it is helpful to ask three meta-questions: first, whether the research plan is feasible given time, access, and resources; second, whether the research questions are ones in which the researcher has a deep and abiding interest and passion; and third, whether the research is worthwhile, ethical, and likely to yield significant new knowledge (i.e. *who benefits*?). Rossman and Rallis (2012: 114–118) call these 'doability' questions, framing them as 'Can do?', 'Want to do?', and

'Should do?'. The answers depend on the researcher's knowledge and prior experiences, subject position, disciplinary orientation, and the social context in which the research takes place. For example, in Meek's (2010) study of language revitalization in a Yukon Kaska community, the research questions emerged not only from scholarship on Indigenous-language endangerment and revitalization, but also from her role as an Indigenous scholar, educator, teaching assistant, and Kaska language learner in that community. Similarly, in long-term ethnographic work in her natal Hopi community in the southwestern US, Nicholas (2014: 75) relates that her interest in studying language shift among Hopi youth came about as she assisted "in providing Hopi literacy lessons to Hopi youth in reservation schools".

Identifying participants and a research setting requires weighing the entailments of one's research questions − the knowledge needed to answer the questions − alongside feasibility, access, and ethics. While ethnographers may choose research settings based on a deliberate weighing of these factors, contexts sometimes 'choose' the researcher based on local needs and qualities inherent in the ethnographer her/him/them self. The optimal research context is often not faraway places or unknown participants, but rather the places and people the researcher already knows well. For example, Blommaert (2013: 52) conducted an ethnographic study of linguistic landscapes in his multilingual inner-city neighborhood in Antwerp, Belgium, where he had lived for 20 years. As part of this "ethnographic immersion", he documented 24 "publicly visible" languages in contact, including Dutch, Turkish, Polish, Thai, Tamil, Serbo-Croatian, and Latin (Blommaert 2013: 52). Similarly, in an ethnographic study arising from "an unplanned set of circumstances" in a remote multilingual Kenyan hillside, Gilmore (2016: xvii) documented the invention of a pidgin, Kisisi ("Our Language"), by two five-year-old boys: her son Colin, an English-speaking American, and his friend Sadiki, a Swahili-speaking Turkana.

As suggested by these examples, in ethnography, the researcher is the primary research 'instrument'. However, the instrument metaphor should not be construed as an antiseptic tool; to the contrary, ethnographic research requires personal involvement over extended periods of time in which the researcher is a learner, an interpreter, and sometimes a collaborator (Rossman and Rallis 2012). This requires careful consideration of both participant-researcher power relations and researcher positionality. Canagarajah (2015: 38) encourages language researchers to ask: "Who are you to your respondents, and what are the effects of your identity on the research?".

Participant observation − engaging appropriately in the day-to-day activities of participants − is the "method of choice" in ethnographic research (Canagarajah 2006: 156). Observations should be systematically recorded in field notes: inscriptions of social life and discourse *in situ* that include a rich description of the setting, participants, and activities (Emerson, Fretz, and Shaw 2011: 12). Field notes also include verbatim quotes of naturally occurring speech and take note of body language and nonverbal communication. In addition, ethnographers may use complementary quantitative methods such as frequency counts to understand the patterning of activity, and demolinguistic mapping of reported home language use across time and space (Extra 2010).

Ethnographic interviews often include casual conversations recorded in the course of participant observation. These unstructured interviews help establish relationships with participants "and create a space within which the participant feels free to tell his or her [or their] story"; this interview format "does not have any 'questions' that are 'asked'

per se but requests information by introducing some very broad topics for discussion within the context of a general conversation" (Olson 2011: 39). Interviews may also be structured, using predetermined, non-variant questions. More common are semi-structured interviews, with a mix of "more and less structured interview questions" carried out one-on-one with key participants and/or in small groups (Merriam and Tisdell 2016: 110). These types of interviews are typically "guided by a list of questions or issues to be explored", allowing the researcher "to respond to the situation at hand, to the emerging worldview of the respondent, and to new ideas on the topic" (Merriam and Tisdell 2016: 110−111). The choice of interview format depends on the research questions, goals, and logistical constraints. Interview data are recorded in field notes and by audiotaping and/or videotaping and are transcribed in written form.

When working in multilingual settings, a key question concerns what language(s) to use in conducting interviews. As Coronel-Molina (2015: 15) describes in his ethnographic study of the High Academy of the Quechua Language in Cuzco, Peru, the decision about whether to conduct interviews in Spanish or Quechua "could have a significant impact on the course of my investigation". Weighing the study's purpose, the likelihood of code-mixing, and the practicalities of participants' linguistic repertoires, Coronel-Molina decided on Spanish. However, in quoting conversations or interviews he included "the speaker's original utterances" in the language in which they occurred "to give more authentic voice to the speakers" (Coronel-Molina 2015: 16). A general rule of thumb is that interviews should be transcribed in the language(s) in which they are conducted, and then translated as needed.

From data gathered through these methods, ethnographic accounts are constructed inductively, "working from empirical evidence towards theory, not the other way around" (Blommaert and Jie 2010: 12). The goal is to arrive at "grounded theories about language as it is practiced in local contexts" (Canagarajah 2006: 153). While building an inductive theory relies on the ethnographer's ability to produce rich description, ethnographically derived interpretations also rely on cross-case and cross-cultural comparison. As Hymes (1980: 90) emphasized, the validity and transferability of ethnographic accounts are greatly enhanced by "contrastive insight" built cumulatively across time and space.

4. Illustrative ethnographies of language contact phenomena

Contemporary ethnographic work on language contact phenomena is shaped by what, at first blush, appear to be contending forces. On the one hand, intensified (trans)migration and globalization are creating 'superdiverse' urban environments (Arnaut et al. 2016). At the same time, these processes intensify pressures for homogenization, widen social and economic stratification, and further marginalize non-dominant languages and communities − forces for language shift that place non-dominant communities (and speakers) at risk and threaten global linguistic and cultural diversity. As Wang et al. (2014: 38) note, globalization affects not only large urban centers, but also nonurban environments − "places not usually […] identified as 'globalized' or 'superdiverse'".

Exemplifying this, King and Haboud (2011) examined the impacts of large-scale mobility on family- and community-based language practices among one Quichua group

in Ecuador, the Saraguros. While much language-related research has focused on those who migrate, these researchers asked how migration affects those "left in its wake": the children and family members who stay behind (King and Haboud 2011: 139). Long-term ethnographic fieldwork revealed that, as parents left agricultural pursuits to seek income abroad, the traditional domains for child language socialization were transformed. As King and Haboud (2011: 148) relate, the time children "once spent working with family around the house and fields is now passed attending school or hanging around town", or in other words, "domains where Spanish dominates". These processes are a force for Quichua language shift.

Canagarajah's (2011) three-year ethnographic study of Sri Lankan Tamils in Toronto, London, and Lancaster, California, provides an intriguing point of comparison to the Saraguro-Quichua case. In all three urban settings, Canagarajah (2011: 79) found "a distinct shift toward English among the younger generation". He attributes this to a complex constellation of sociohistorical, sociolinguistic, and sociocultural factors: the legacy of British colonization and "a deeply engrained attitude" of English superiority, the privileging of English in education and employment, and community members' ability to engage in Sri Lankan Tamil cultural practices despite limited Tamil speaking ability (Canagarajah 2011: 92–93). This study highlights the value of ethnography in illuminating the local, context-specific challenges of sustaining heritage languages in contact with dominant languages in a globalizing world system. As Canagarajah (2011: 78) points out, "[d]iaspora communities negotiate not only home and migrant locations, but also lateral relationships with their people in other migrant locations to construct a virtual community and more hybrid identities with unpredictable linguistic possibilities".

These are a few examples of the ways in which ethnographic inquiry sheds light on how the "infrastructures of globalization" (Wang et al. 2014: 38) impact language choices, practices, and futures in diverse zones of linguistic contact. Ethnographers have also investigated the ways in which larger policy structures and globalizing forces are reconfigured in local language ideologies and practices, language contact phenomena that Hornberger and McCarty (2012) theorize as "globalization from the bottom up". In an ethnographic analysis of bilingual education in Mozambique, Chimbutane and Benson (2012) show how local appropriations of top-down curricular reforms open new spaces for the promotion of Indigenous languages and cultures. Working in a South African undergraduate language program, Joseph and Ramani (2012) show how a focus on additive multilingualism in teacher preparation can unseat the hegemony of English within the 'new globalism'. In these and other ethnographic cases, relatively small-scale language policy reformulations create new options through which non-dominant languages historically constructed as "traditional" (and hence as 'not useful' in the global economy) can be repositioned and re-signified as "modern" (Joseph and Ramani 2012: 32).

Another important stream of ethnographic work has explored the interactions among speakers within particular linguistic ecologies, a concept introduced in 1972 by Einar Haugen. As Hornberger and Hult (2008: 281) explain, the "ecology of language has […] developed as a conceptual approach to investigating how linguistic ecologies relate to social, historical, sociolinguistic, and political forces at individual, community, and societal levels of social organization". Much ethnographic work within this conceptual framing has centered on education. Studies in Hornberger (2003a), for instance, examine language ecologies and ideologies in multilingual classrooms in diverse and dispersed village pri-

mary schools on Borneo and Corsica, urban primary schools in the US and the UK, and secondary public and community schools in Liverpool, London, and Göteborg. Based on rich, in-depth ethnographic research, these studies offer emic or 'insider' views of language contact through detailed descriptions and analysis of verbatim transcriptions of multilingual classroom interactions.

Hult (2003) conducted an ethnographic ecolinguistic study of the complex ways in which English is integrated with Swedish in daily life in Lund and Malmö, Sweden. Hult's (2003: 51) fieldwork took place over three months, during which data were collected using "photography, field observations/notes, and informal interviewing of a variety of people [as well as] online archives of the Swedish newspapers *Svenska Dagbladet* and *Metro* (Skåne)". Hult shows that English remains an articulated component of Swedish life, and that both languages have their place in Swedish society. One implication of this study is that an "overt language policy […] about Swedish linguistic culture may be a welcome codification of certain current linguistic practices while also serving to ensure that everyone receives access to the linguistic resources required" (Hult 2003: 61).

Wyman (2012) also used a linguistic ecological approach in a multiyear ethnographic study of youth peer culture and language shift in Piniq, a Yup'ik Eskimo village in Alaska. Drawing on participant observation, and formal and informal interviews with parents, educators, and youth, Wyman (2012: 2) argues that in dynamic situations of language shift such as Piniq, it is essential to examine "how language learning and beliefs about languages change over time within complex *linguistic ecologies* − social networks of language use that root within, and extend beyond specific communities and geographical places". Wyman (2012: 272) places schooling squarely within these linguistic-ecological considerations, showing how the local school was "undermining heritage language maintenance", but also revealing its potential to support youths' bilingualism through consistent, community-driven bilingual education programming.

Linguistic landscape studies (LLS) represent another stream of language contact research. LLS focus on "language in the environment, words and images displayed and exposed in public spaces" (Shohamy and Gorter 2009: 1). Much of the empirical work on linguistic landscapes has been conducted using digital camera technology and text analysis (Gorter 2006). However, precisely because of the linguistic diversity it seeks to document and understand, ethnography is uniquely positioned to capture the social meaning of "situated signs-in-public-space", and to identify "the fine fabric of their structure and function in constant interaction with several layers of context" (Blommaert 2013: 20).

In illustration of this, Curtin (2009) conducted a critical social semiotic linguistic landscape study of indexical signs and identities in Taipei, Taiwan. Using ethnographic methods and discourse analysis of non-Chinese scripts from four major sections of the city, including permanent street and building signage, and salient displays of language on vehicles, posters, and windows, she developed "preliminary (*etic*) categories based on source language and the apparent function and meaning of each" (Curtin 2009: 225). She then worked for 10 months with local Taiwanese with varying language backgrounds (Mandarin-only, Mandarin/Southern Min, Southern Min/Mandarin) to develop "a more *emic* framework" incorporating characteristics of interpretive readings (Curtin 2009: 225). Curtin (2009: 233) concludes that, "the graphic regime of orthographic scripts is

an important part of the ongoing process of identification, a process that is historically situated and ideologically framed in a dialogic relationship to the Other".

Malinowski (2009: 107) undertook an ethnographic study of linguistic landscape authorship in California, exploring "the symbolic and political significance of a particular linguistic code's appearance with other codes in bilingual signs". This four-month study included interviews, participant observation, photograph and media analysis, and interpretive walking and driving tours. Malinowski's (2009: 124) study indicates that "seemingly intentional meanings [of signage] can in fact remain hidden to the writers of signs, arising instead from larger historical processes that have become sedimented into practices of literacy and technologies of design". Further, he demonstrates that there must be a distinction between the multifaceted, distributed process of linguistic landscape authorship, and the individual social actors.

5. Challenges and opportunities

A prime challenge in ethnographic research emerges precisely from what makes it so valuable in understanding complex language contact phenomena: ethnography is a time-intensive form of inquiry. This means that there is often a significant lag between what is observed (ethnography as process) and what is reported (ethnography as product). Yet, even with extended time commitments, ethnographers cannot attend to every detail in fieldwork, nor is this desirable. What we choose to focus on and edit out is steeped in our disciplinary training, and in our multifaceted subject position as distinctly raced, classed, gendered, and abled, and as both the observer-inquirer and the author of the account. Like other researchers, ethnographers must constantly exercise reflexivity − the critical evaluation of one's own subject position and the power relations that influence the research. This methodological positioning aligns with a growing body of work in critical ethnography, which interrogates ethnographic authority and the researcher's stance and positionality, recognizing that all knowledge production "is implicated in power relations" (Canagarajah 2006: 156; see also Canagarajah 2015; Lin 2015).

Perhaps a more consequential challenge in ethnographically oriented language contact research is such work's potential to transform the linguistic inequalities it exposes. We view this as both a challenge and an opportunity, as ethnographers balance a critical understanding of linguistic inequalities with an understanding of the ability of stakeholders − both researchers and participants − to employ that understanding in actionable and anti-hegemonic ways.

From early ethnographers' attention to firsthand fieldwork, to the ethnography of communication, to recent work addressing the complex realities of globalization, superdiversity, linguistic landscapes, and language shift and reclamation, ethnography affords rich, multilayered insights into the nature and meanings of language processes in particular sociocultural settings. These insights stem from a distinctive 'way of seeing' language contact phenomena through a critical cultural lens, and a 'way of looking' at those processes up close, over an extended period of time, and from participants' point of view. By casting an "ethnographic eye" on language contact phenomena, researchers can "uncover the indistinct voices, covert motivations, embedded ideologies, invisible instances, or unintended consequences" of those phenomena as they are manifested in everyday life (Hornberger and Johnson 2007: 24).

Moreover, as a form of knowledge production, ethnography is intrinsically democratizing, as its primary goal − understanding "the meanings, norms, and patterns of a way of life" (Hymes 1980: 98) − reflects what people do in everyday situations of language contact. Ethnography thus has the potential to break down hierarchies between the 'knower' and the 'known', and to bring local stakeholders directly into the research process. This 'way of being' a researcher unsettles the notion of value-free research, replacing it with grounded forms of critical, collaborative, reflexive inquiry. This may be ethnography's most important contribution to the study of language contact phenomena: taking seriously the need for collaboration and reciprocity with research participants in the quest for sociolinguistic knowledge and social justice.

Acknowledgments

Parts of this chapter are updated from McCarty (2015a, 2015b) and McCarty and Liu (2016). The authors wish to thank the editors, copyeditor, and two anonymous reviewers for their very helpful suggestions on earlier versions of this chapter.

6. References

Anderson-Levitt, Kathryn
2006 Ethnography. In Judith L. Green, Gregory Camilli & Patricia B. Elmore (eds.), *Handbook of complementary methods in education research*, 279−295. Washington, DC: American Educational Research Association.

Androutsopoulos, Jannis (ed.)
2006 Introduction: Sociolinguistics and computer-mediated communication. *Journal of Sociolinguistics* 10(4). 419−438.

Arnaut, Karel, Jan Blommaert, Ben Rampton & Massimiliano Spotti (eds.)
2016 *Language and superdiversity*. New York & London: Routledge.

Baker, Lee D.
1998 *From savage to Negro: Anthropology and the construction of race, 1896−1954*. Berkeley: University of California Press.

Bigelow, Martha & Johanna Ennser-Kananen (eds.)
2015 *The Routledge handbook of educational linguistics*. New York & London: Routledge.

Blommaert, Jan
2013 *Ethnography, superdiversity, and linguistic landscapes: Chronicles of complexity*. Bristol, UK: Multilingual Matters.

Blommaert, Jan & Dong Jie
2010 *Ethnographic fieldwork: A beginner's guide*. Bristol, UK: Multilingual Matters.

Canagarajah, Suresh
2006 Ethnographic methods in language policy. In Thomas Ricento (ed.), *An introduction to language policy: Theory and method*, 153−169. Malden, MA: Blackwell.

Canagarajah, A. Suresh
2011 Diaspora communities, language maintenance, and policy dilemmas. In Teresa L. McCarty (ed.), *Ethnography and language policy*, 77−97. New York & London: Routledge.

Canagarajah, A. Suresh
2015 Ethical considerations in language policy research. In Francis M. Hult & David Cassels Johnson (eds.), *Research methods in language policy and planning*, 33−44. Malden, MA: Wiley Blackwell.

Chimbutane, Feliciano & Carol Benson
2012 Expanded spaces for Mozambican languages in primary education: Where bottom-up meets top-down. *International Multilingual Research Journal* 6(1). 8–21.
Clyne, Michael
2003 *Dynamics of language contact*. Cambridge, UK: Cambridge University Press.
Combs, Mary Carol, Norma González & Luis Moll
2011 US Latinos and the learning of English: The metonymy of language policy. In Teresa L. McCarty (ed.), *Ethnography and language policy*, 185–203. New York & London: Routledge.
Coronel-Molina, Serafín M.
2015 *Language ideology, policy, and planning in Peru*. Bristol, UK: Multilingual Matters.
Curtin, Melissa L.
2009 Languages on display: Indexical signs, identities, and the linguistic landscape of Taipai. In Elana Shohamy & Durk Gorter (eds.), *Linguistic landscape: Expanding the scenery*, 221–237. New York & London: Routledge.
Emerson, Robert M., Rachel I. Fretz & Linda L. Shaw
2011 *Writing ethnographic fieldnotes*, 2nd edn. Chicago & London: University of Chicago Press.
Erickson, Frederick
2011 A history of qualitative inquiry in social and educational research. In Norman K. Denzin & Yvonna S. Lincoln (eds.), *The SAGE handbook of qualitative research*, 43–59. Los Angeles, CA: SAGE.
Extra, Guus
2010 Mapping linguistic diversity in multicultural contexts: Demolinguistic perspectives. In Joshua A. Fishman & Ofelia García (eds.), *Handbook of language and identity: Disciplinary and regional perspectives*, vol. 1, 2nd edn., 107–122. New York: Oxford University Press.
García, Ofelia & Li Wei
2014 *Translanguaging: Language, bilingualism, and education*. London, UK: Palgrave Macmillan.
Geertz, Clifford
1973 *The interpretation of cultures*. New York, NY: Basic Books.
Gilmore, Perry
2016 *Kisisi (Our Language): The story of Colin and Sadiki*. Malden, MA: Wiley Blackwell.
González, Norma
2001 *I am my language: Discourses of women and children in the borderlands*. Tucson: University of Arizona Press.
Gorter, Durk (ed.)
2006 *Linguistic landscape: A new approach to multilingualism*. Clevedon, UK: Multilingual Matters.
Haugen, Einar
1972 *The ecology of language: Essays by Einar Haugen*. Stanford, CA: Stanford University Press.
Heath, Shirley Brice
1983 *Ways with words: Language, life, and work in communities and classrooms*. Cambridge, UK: Cambridge University Press.
Hinton, Leanne, Leena Huss & Gerald Roche (eds.)
2018 *The Routledge handbook of language revitalization*. New York & London: Routledge.
Hornberger, Nancy H. (ed.)
2003a *Continua of biliteracy: An ecological framework for educational policy, research, and practice in multilingual settings.* Clevedon, UK: Multilingual Matters.

Hornberger, Nancy H.
2003b Linguistic anthropology of education (LAE) in context. In Stanton Wortham & Betsy Rymes (eds.), *Linguistic anthropology of education*, 245–270. Westport, CT: Praeger.
Hornberger, Nancy H. & David Cassels Johnson
2007 Slicing the onion ethnographically: Layers and spaces in multilingual language education policy and practice. *TESOL Quarterly* 41(3). 509–532.
Hornberger, Nancy H. & Francis M. Hult
2008 Ecological language policy. In Bernard Spolsky & Francis M. Hult (eds.), *The handbook of educational linguistics*, 280–296. Malden, MA: Blackwell.
Hornberger, Nancy H. & Teresa L. McCarty (eds.)
2012 Globalization from the bottom up: Indigenous language planning and policy across time, space, and place. [Special issue]. *International Multilingual Research Journal* 6(1).
Hornberger, Nancy H., Aldo Anzures Tapia, David H. Hanks, France Kvietok Dueñas & Siwon Lee
2018 Ethnography of language planning and policy. *Language Teaching* 51(2). 152–186.
Hult, Francis
2003 English on the streets of Sweden: An ecolinguistic view of two cities and a language policy. *Working Papers in Educational Linguistics* 19(1). 43–63.
Hymes, Dell
1964 Introduction: Toward ethnographies of communication. *American Anthropologist* 66(6). 1–34.
Hymes, Dell
1974 The use of anthropology: Critical, political, personal. In Dell Hymes (ed.), *Reinventing anthropology*, 3–79. New York, NY: Vintage Books.
Hymes, Dell
1980 *Language in education: Ethnolinguistic essays*. Washington, DC: Center for Applied Linguistics.
Hymes, Dell
1996 *Ethnography, linguistics, narrative inequality: Toward an understanding of voice*. London: Taylor & Francis.
Johnson, David Cassels (ed.)
2013 Ethnography of language policy. [Special issue]. *International Journal of the Sociology of Language* 219.
Joseph, Michael & Esther Ramani
2012 "Glocalization": Going beyond the dichotomy of global versus local through additive multilingualism. *International Multilingual Research Journal* 6(1). 22–34.
King, Kendall A. & Marleen Haboud
2011 International migration and Quichua language shift in the Ecuadorian Andes. In Teresa L. McCarty (ed.), *Ethnography and language policy*, 139–159. New York & London: Routledge.
Leonard, Wesley Y. & Haley De Korne (eds.)
2017 Reclaiming languages: Contesting and decolonising 'language endangerment' from the ground up. [Special issue]. *Language Documentation and Description* 14.
Lim, Lisa & Umberto Ansaldo
2016 *Languages in contact*. Cambridge, UK: Cambridge University Press.
Lin, Angel M. Y.
2015 Researcher positionality. In Francis M. Hult & David Cassels Johnson (eds.), *Research methods in language policy and planning*, 21–32. Malden, MA: Wiley Blackwell.
Malinowski, David
2009 Authorship in the linguistic landscape: A multimodal-performative view. In Elana Shohamy & Durk Gorter (eds.), *Linguistic landscape: Expanding the scenery*, 107–125. New York & London: Routledge.

Martin-Jones, Marilyn, Adrian Blackledge & Angela Creese (eds.)
2012 *The Routledge handbook of multilingualism*. London & New York: Routledge.
McCarty, Teresa L. (ed.)
2011 *Ethnography and language policy*. New York & London: Routledge.
McCarty, Teresa L.
2015a Ethnography in educational linguistics. In Martha Bigelow & Johanna Ennser-Kananen (eds.), *The Routledge handbook of educational linguistics*, 23–37. New York & London: Routledge.
McCarty, Teresa L.
2015b Ethnography in language planning and policy research. In Francis M. Hult & David Cassels Johnson (eds.), *Research methods in language policy and planning: A practical guide*, 81–93. Malden, MA: Wiley Blackwell.
McCarty, Teresa L. & Lu Liu
2016 Ethnography of language policy. In Kendall A. King, Yi-Ju Lai & Stephen May (eds.), *Encyclopedia of language and education: Research methods*, vol. 10, 3rd edn., 53–66. Cham, Switzerland: Springer International.
Meek, Barbra A.
2010 *We are our language: An ethnography of language revitalization in a Northern Athabaskan community*. Tucson: University of Arizona Press.
Merriam, Sharon B. & Elizabeth J. Tisdell
2016 *Qualitative research: A guide to design and implementation*, 4th edn. San Francisco, CA: Jossey-Bass.
Nicholas, Sheilah E.
2014 "Being" Hopi by "living" Hopi – Redefining and reasserting cultural and linguistic identity: Emergent Hopi youth ideologies. In Leisy T. Wyman, Teresa L. McCarty & Sheilah E. Nicholas (eds.), *Indigenous youth and multilingualism: Language identity, ideology, and practice in dynamic cultural worlds*, 70–89. New York & London: Routledge.
Olson, Karin
2011 *Essentials of qualitative interviewing*. Walnut Creek, CA: Left Coast Press.
Onysko, Alexander (ed.)
2016 Language contact and World Englishes. [Special issue]. *World Englishes* 35(2).
O'Rourke, Bernadette & John Walsh (eds.)
2018 Comparing "new speakers" across language contexts: Mobility and motivations. [Special issue]. *Journal of Multilingual and Multicultural Development* 39(5).
Paris, Django & Maisha T. Winn (eds.)
2014 *Humanizing research: Decolonizing qualitative inquiry with youth and communities*. Los Angeles, CA: SAGE.
Philips, Susan U.
1983 *The invisible culture: Communication in classroom and community on the Warm Springs Indian reservation*. New York & London: Longman.
Pike, Kenneth
1967 *Language in relation to a unified theory of the structure of human behavior*, 2nd edn. The Hague: Mouton.
Romero-Little, Mary Eunice, Teresa L. McCarty, Larisa Warhol & Ofelia Zepeda
2007 Language policies in practice: Preliminary findings from a large-scale national study of Native American language shift. *TESOL Quarterly* 41(3). 607–618.
Rossman, Gretchen B. & Sharon F. Rallis
2012 *Learning in the field: An introduction to qualitative research*, 3rd edn. Los Angeles, CA: SAGE.
Shohamy, Elana & Durk Gorter (eds.)
2009 *Linguistic landscape: Expanding the scenery*. New York & London: Routledge.

Smith, Linda Tuhiwai
2012 *Decolonizing methodologies: Research and Indigenous peoples*, 2nd edn. New York: Zed Books.
Spolsky, Bernard & Francis Hult (eds.)
2008 *The handbook of educational linguistics*. Malden, MA: Blackwell.
Tollefson, James W.
2013 Language policy in a time of crisis and transformation. In James W. Tollefson (ed.), *Language policies in education: Critical issues*, 11–34. New York & London: Routlege.
Wang, Xuan, Massimiliano Spotti, Kasper Juffermans, Leonie Cornips, Sjaak Kroon & Jan Blommaert
2014 Globalization in the margins: Toward a re-evaluation of language and mobility. *Applied Linguistics Review* 5(1). 23–44.
Wolcott, Harry
2008 *Ethnography: A way of seeing*, 2nd edn. Lanham, MD: AltaMira Press.
Wyman, Leisy T.
2012 *Youth culture, language endangerment, and linguistic survivance*. Bristol, UK: Multilingual Matters.

Teresa L. McCarty, Los Angeles (USA)
Lu Liu, Los Angeles (USA)

49. Interviewing

1. Introduction
2. Contemporary interviewing
3. Collecting data
4. Critiques of the SLI
5. Interactional roles and goals
6. Fieldworker profiles and approach
7. Conclusions
8. References

1. Introduction

Both in linguistics and the social sciences at large, the interview has been an important research tool for collecting data since the beginning of the 20th century (Chambers 2008; Platt 2012). It was and remains a key method in fieldwork strategies both to yield opinions about language and to prompt different forms and styles of speech. The primary goal of interviewing is to elicit the linguistic and social information we require to answer our questions about languages and the sociocultural context(s) in which they are spoken. Interviewing in language contact communities offers opportunities to explore intricate linguistic and social relationships. The fieldwork process itself can contribute to our understanding of the communities with which we engage and the structures and functions of their language(s).

This chapter discusses issues and practices relevant to interviewing. It provides a brief history of the interview in sociolinguistic research – including critique – and describes goals and methods of interviewing, paying particular attention to language con-

https://doi.org/10.1515/9783110435351-049

tact scenarios. Although language contact research requires heightened sensitivity to social and linguistic complexity, the theoretical and methodological approaches to data collection through interviewing are consistent with those in any linguistic community.

1.1. History of the interview

Contemporary sociolinguistic interviewing arguably owes its roots to methods established by dialect geographers in the late nineteenth and the first half of the twentieth centuries. In projects such as the *Atlas Linguistique de la France* (Gilliéron and Edmont 1969), the *Atlante italo-svizzero* (Jaberg, Jud, and Scheuermeier 1928), and the *Atlas Lingüístico de la Península Ibérica* (Navarro Tomás 1933−1962; García Mouton et al. 2016) in Europe, as well as large-scale dialect atlases in North America, fieldworkers were charged with eliciting lexical items as well as specific phonetic and grammatical structures. Their methods, interviews of sorts, included asking participants direct and indirect questions that were expected to prompt responses with certain phonetic or grammatical variables or lexical items. For example, particular lexical items were shown in photos or referred to indirectly in an effort to investigate regional lexical and phonetic variation. Current interviewing makes use of comparable approaches in the form of word lists or picture tasks.

In this early work, specific methods for elicitation varied according to the interviewer and the project. Fieldworkers were trained in some type of phonetic transcription, and their transcriptions and notes served as the source of data. Fieldworkers also collected demographic information from participants, sometimes involving them in conversations and even participating in ethnographic work. Arguably, these dialectologists were interviewers who engaged participants in thematic conversations to elicit responses beyond the strict purview of their surveys (Johnstone 2000: 8), although in most cases, only their phonetic transcriptions and notes pertinent to the specific goals of the projects remain.

2. Contemporary interviewing

2.1. The sociolinguistic interview

The middle of the 20th century marked the advent of modern sociolinguistics and the recorded interview as one of its prime methodological tools. With advances in technology, it became possible to record speech more easily using portable devices. Labov introduced methods for recording speakers for research on language variation and change (see Tagliamonte 2016). Although refined over time, many of the principles first raised in his 1963 study of phonetic variation on Martha's Vineyard remain constant. Labov (1963: 12−13) underscores the importance of eliciting sufficient instances of targeted linguistic variables in different speech styles. He describes these styles (casual, emotional, careful, and reading), and outlines techniques used to encourage such speech, including general questions, questions prompting responses with specific words, and a passage and lexical questionnaire with instances of the features to be read by participants.

Methods from this study formed the basis of the most common method of collecting recorded speech in sociolinguistics, the Sociolinguistic Interview (SLI). Gumperz's (1964) contemporary fieldwork in Hemnes offered a slightly different, complementary approach, focusing on field observations and subsequent recorded interviews of groups of local speakers (rather than individual interviews) to investigate code switching, a contact phenomenon. The effect of the researcher's presence (or the 'observer's paradox') was important to both scholars, who sought strategies to minimize it. Labov's (1966, 1972) subsequent work refined techniques for interviewing within a variationist framework (including the deliberate use of group interviews). Labov (1984) explicitly outlines these methods, incorporating guides for conversation-prompting question modules on thematic topics and the elicitation of more careful styles through texts read aloud, which are discussed further below, along with other perspectives.

2.1.1. Integrated approaches: interviewing and ethnography

As I emphasize throughout this chapter, an understanding of communities is key to successful interviewing, and ultimately, to the analysis and interpretation of the data collected in interviews. Familiarizing oneself with social context, and the use and role of language(s), is a part of the research process. Indeed, some insights may come in hindsight upon reconsidering and re-contextualizing the fieldwork process and the data collected (Mendoza-Denton 2008: 38–40). Interviewing in language contact situations can involve differing degrees of integration in communities. Ethnographic research is a gold standard for any sociolinguistic work, offering invaluable insights in local meaning. It is through immersion in a community as a participant-observer that the most nuanced understandings of language in its social context can be gleaned. That researchers must be "physically and socially present" is, of course, a challenge on several levels (Besnier and Philips 2014: 128). This type of fieldwork involves a substantial commitment of time and personal energy, and researchers may not be able to participate to the same degree in communities of interest because of their social or linguistic profile. Nevertheless, the rewards of an ethnographic approach are substantial. Mendoza-Denton's (2008) oft-cited work with Latina gang girls is the leading example of ethnographic work with a variationist perspective in a language contact scenario. Although not focusing on language contact per se, it serves as a model for insightful analyses of linguistic features and language use in a context in which Spanish and English and their influences carried socio-symbolic meaning. Mendoza-Denton's background afforded her a credible yet liminal position from which to engage with Latina youth gangs, and her work in a tutoring center facilitated involvement with the youth. Both qualitative observations and data collection allowing for quantitative analyses were essential to this work, and interviewing in both Spanish and English was part of a comprehensive repertoire of connection and data collection.

2.1.2. Interviewing outside of ethnography

Researchers lacking the resources to enmesh themselves for extensive periods of time in ethnographic work may still engage successfully with a community, and supplement

interviews by researching its social profile and connecting with community members in informal contexts. In considering critiques of data collection through the sociolinguistic interview, Meyerhoff (2016: 439) states that "it's not a question of whether interview or ethnographic approaches are a priori 'better'; instead, researchers need to be clear about what their questions are and what methods will best address those questions".

Approaches to interviewing derive from the goals of individual research projects. In language contact research, investigating phenomena such as lexical borrowings, potential structural transfer, code-mixing, or attitudes toward language use and shift may call for adapting the methods of the traditional SLI. As Becker (2018: 101) observes, the SLI was designed "to provide evidence of sociolinguistic variation". Over time, the term has come to be used beyond its original and specific intention, and is commonly applied to any recorded interview in sociolinguistic work, even those which do not necessarily follow Labov's specific methods and definitions (Becker 2018; Meyerhoff 2016). Many principles from which the SLI arose are relevant to broader work, and adapting interview methods does not have to be problematic, provided researchers' methods are clearly described and replicable.

3. Collecting data

3.1. Social factors and qualitative interviewing

This chapter's focus on data collection within a quantitative paradigm should not be interpreted as negating the value of qualitative approaches. Indeed, interviews are an important source of both qualitative and quantitative data, and the two are not mutually exclusive. In research focusing on social aspects of language like language use and language attitudes, interviews are often more structured, with rosters of specific questions or full questionnaires for quantitative analyses (Holmquist and Kahn 2017).

This is entirely appropriate, as set questions are advantageous for comparability in data analysis. Such protocols can involve questions which are posed directly and indirectly, and invite closed and open-ended responses. Even in the context of more rigid protocols, participant-initiated conversation around relevant themes can be valuable, both informing and enriching the original research questions. For example, Guardado and Becker (2014) describe qualitative methods used to gain insights into the status of Spanish among Spanish-speaking families living in English-dominant Vancouver, Canada. Guardado (2002) used semi-structured interviews to ask participating families (either as individuals or groups) about their language practice and history, both pursuing targeted lines of inquiry and allowing themes to emerge from their narratives. This was coupled with hours of participant observation, allowing him to triangulate data for the best analyses. Sharma (2011) included two sets of interviews as part of a multifaceted fieldwork strategy, with one devoted to the social context of speakers and their language use. Ultimately, qualitative and quantitative analyses are complementary, helping us to glean valuable emic information about communities and their languages, while also contributing to a more nuanced interpretation of larger amounts of data.

3.2. Quantitative data

Quantitative research on the linguistic consequences of language contact requires sufficient recorded speech for analysis. This typically entails interviews of one and a half to two hours of conversational speech. Naturalistic, unselfconscious, spontaneous speech, preferably in the form of personal narratives, has been the sought-after ideal in interviewing (Schilling 2013). Consequently, the biggest challenge in interviewing is mitigating the nature of the interview itself.

Variationist work has long used the term *vernacular* to refer to the least self-conscious, most naturalistic speech style. I am careful not to use it here, since, as Becker (2018) and Meyerhoff (2016) discuss, the vernacular as a construct can be fraught with challenges: associated with higher rates of nonstandard features or idealized for its purported authenticity. The potential for essentializing the vernacular is problematic (Meyerhoff 2016: 440), and it is important to recognize and validate many types of naturalistic (and indeed, natural) speech

Recording speech for the sole purpose of academic study invites artificiality, or at the very least formality, and researchers are faced with the aforementioned 'observer's paradox': wanting to observe the way people speak when they are not being observed (Labov 1984: 30). Although it is impossible to avoid this dilemma altogether, there are strategies to diminish its effect. Fundamentally, they involve minimizing awkwardness inherent to a recorded interview due both to technical aspects (the presence of recording equipment) and to the relationship of the interviewer with consultant(s). The connections and relationships between fieldwork interviewers and speaker(s) will be discussed further below. Although I generally focus on interview techniques for analyzing linguistic variation, these principles apply to interviewing in general.

3.3. Eliciting spontaneous speech

In order to encourage naturalistic speech, researchers typically prepare an Interview Schedule: questions related to a central theme organized in modules that are introduced by interviewers both to elicit speech and to gain demographic and cultural information from participants (Labov 1984: 33). These modules normally address interrelated aspects of the speaker's experience, such as family, memories of childhood, cultural traditions, and language use. Often, a module begins with a more general question, and speakers are encouraged to elaborate with a personal narrative related to the topic. While this practice can guide the interview, topics introduced by consultants are definitely encouraged and sometimes preferred. For example, in Poplack's (2017) work on loanwords, she describes interviewing for her Ottawa-Hull corpus of French speakers with differing degrees of contact with English. Although (community insider) fieldworkers had prepared question modules, consultants largely initiated topics, enabling collection of spontaneous, informal speech styles.

Eliciting naturalistic speech requires a fine balance between speaking and listening. It is appropriate for interviewers to participate in the conversation but also essential that they neither dominate nor interrupt, since the primary goal is to record as much speech as possible from the participant. The phrasing of questions is also important, avoiding 'yes/no'-type questions and instead using open-ended questions that encourage longer

responses. Asking exploratory questions at the beginning of an interview can also be useful for assessing interviewees' interests to use them as a springboard for conversation (Tagliamonte 2006: 39). Ultimately, flexibility and responsiveness are paramount to success. An effective interview should have the tone of an informal conversation and flow easily from topic to topic. The best interviews become speaker driven as they themselves introduce topics and share their experiences, obviating the need to refer to banks of questions. Consultants should feel as comfortable as possible and unpressured about their opinions. This is easier further along into fieldwork, as interviewers learn more about communities and can build on successful topics from prior interviews.

3.4. Beyond conversation: targeting features

Labov's (1972) early work and innumerable subsequent studies have made use of word lists and reading passages designed both to target phonetic features under investigation and to address speech styles. These tools are relevant to contemporary studies, despite shifts in approaches to style. Reading styles may still be formal and serve to gauge speakers' attitudes toward prestige. Additionally, some linguistic features, including those relevant to contact research, are not as common in conversational data. For example, phonological variables may occur only rarely in phonetic contexts conducive to acoustic analyses. Directing conversation to elicit particular words in these contexts may result in frustration or at best, stilted conversation. In these cases, word lists, lists of sentences or longer reading passages allow for analyses of otherwise unavailable features.

In language contact scenarios, it is important to consider both speakers' proficiency in and access to differing styles. For example, in an immigrant minority language community, second or third generation speakers may not read or write their heritage language. Thus, reading components may not elicit the desired data and may make participants feel uncomfortable, potentially compromising the interview somewhat. Researchers may tailor texts to the community rather than eliminating this component of an interview altogether. Mesthrie (2012: 378) reports that his choice to use only a word list, but no reading passage, as part of an interview protocol was based on his awareness that his South African participants may have had difficulty reading, particularly in a language other than their home language. In my own research on Spanish-English bilingual youth in Toronto (Hoffman 2010), I encountered this challenge: while my consultants were high school students or graduates literate in English, some had difficulty reading the Spanish passage that I had assumed was relatively simple.

Interviewers in language contact scenarios may benefit from including components beyond those of the traditional SLI to elicit borrowings and/or potentially contact-induced features. Eliciting well-known stories or recipes that call for the use of structures or lexical items or asking for directions to local sights (a potential adaption of a map task) may help target features. Knowledge of the community will help adapt aspects of the interview. In their work on the use of Spanish and Kaqchikel, a Mayan language, Holmquist and Kahn (2017: 16) relied on community member research assistants to develop assessments of proficiency in Kaqchikel: one assessment targeted speakers' range of vocabulary and their ability to recognize more basic terms, either by translating

the words or by explaining their meaning indirectly. Another component involved responding appropriately (either in Kaqchikel or in Spanish) to basic questions.

3.5. Investigating style

Eliciting differing speech styles can be complicated, as there are many ways to conceptualize 'style' in contact situations as well as in monolingual contexts. Speakers are agentive and performative, orienting themselves toward various aspects of identity in conversation. Work focusing on these aspects of style, including stance and interactional aspects of speech, tends to employ more ethnographic approaches to collect data and to analyze its social meaning(s) (Podesva 2007). Traditional variationist approaches to style have invoked attention to speech to differentiate between more and less careful styles (making use of the aforementioned reading passages, sentences, and word lists) and a range of conversational styles. In contact scenarios, languages may be mixed strategically to present styles; indeed, code-switching (see Chapter 14) can represent an act of identity. Given this added layer of complexity, interviewers must be mindful of the nature of style within the social context of the speaker (and of the interview). Mixing may be a common, unmarked way of speaking for communities. In projects investigating mixed speech (Torres Cacoullos and Travis 2016, 2018) insiders fluent in community norms may be the most suitable interviewers since interviewers and consultants tend to align their switching and language use (Valdés Kroff and Fernández-Duque 2017: 226). Further, the parameters of style can differ in diasporic or obsolescing minority language communities. Because speakers may only use their language in informal contexts, they may not be familiar with the range of styles found in their countries of origin. Elicitation of traditionally formal styles (or the range of styles available in majority language contexts) may not be possible at all.

Fieldworkers' social profiles can influence both style (Bell 1984) and social information in interviews. Bell and Johnson's (1997) study of discourse variables and style in a language contact scenario between the English spoken by Maori and Pakeha (European-descent) Zealanders from similar socioeconomic backgrounds presents the effect of different interviewers in this context. Their findings reflect not only differences between interviewers, recalling Bell's (1984) Audience Design, but also differences in topic or referee design, i.e. the use of features associated with ethnicity to mark that ethnicity in a discussion of cultural practice(s).

3.6. Individual and group interviews

The number of consultants present is also relevant to questions of style and elicitation of spontaneous speech. Although sociolinguistic interviews with individual speakers are most common, interviews with pairs and groups of speakers can be an effective tool for both qualitative and quantitative work. Earlier (Gumperz 1964; Labov 1972) as well as more recent (Cheshire et al. 2011) research has made use of recorded interviews with groups of peers as a strategy to elicit more naturalistic speech. The dynamic of a group interview may minimize the effects of observation, as the interviewer is one of several

speakers, who ideally converse more readily with their peers. Group or pair interviews may be particularly advantageous in language contact situations, as may including local community members in interviews with outsider researchers (Waltermire 2014), encouraging the emergence of more wide-ranging code-mixing practices and local features. However, group interviews may not be possible where there are few remaining speakers or a single speaker of a language (Sabino 2012).

4. Critiques of the SLI

There have been a number of methodological assessments of the sociolinguistic interview. Wolfson (1976) offers an early critique of its utility as a method for the collection of naturalistic speech. At the crux of her argument is the nature of the interview as a recognizable speech event (Hymes 1974), with norms and expectations familiar to the participants. The rules of an interview involve the asking and answering of questions, which precludes in some sense the type of natural speech researchers seek. She distinguishes between a formal interview and a more spontaneous interview; in the latter, fieldworkers endeavor to break with a strict question-answer format and encourage participants to initiate topics and engage in an informal conversation (as described above). According to Wolfson (1976: 195), this strategy brings its own challenges, as the purpose of the encounter is mysterious and ambiguous to the speaker. These arguments are valid. However, since recorded speech is necessarily a primary source of sociolinguistic data, we must recognize them and seek to minimize any inherent awkwardness by using different techniques to elicit the type of speech and information we seek. These concerns also interact with the dynamics within the interviews and the roles of and relationships between interviewers and interviewees. For recent thoughtful discussions contextualizing the SLI and variationist methods, see Becker (2018) and Meyerhoff (2016).

Briggs (1986) underscores the value of understanding how local ways of communicating cultural norms can influence the content and outcomes of interviews. One of his key points addresses sensitivity to ways of asking: framing questions appropriately and bearing in mind the relative intentions of the interviewers and consultants. Ways of asking and raising relevant topics may differ across communities, and familiarity with these practices can contribute to the type of information that participants share.

On a fundamental level, both fieldworkers and research participants are aware that interviews take place to fulfill the needs of a research project. In that sense, they are never entirely socially spontaneous or natural. Ethical considerations for human participants in research projects also constrain the collection of recorded speech. Consultants must give oral or written consent in order for their speech to be recorded, contributing to the sense of formality in an interview context. Furthermore, in most cases, researchers must share information about the nature of the project with them. Although this may complicate the goals of the interview somewhat, it does not necessarily compromise them altogether. General descriptions suffice in most cases, so it is not necessary to reveal the researchers' exact focus or the linguistic variables under investigation. Moreover, sharing some context with speakers may even help address one of Wolfson's critiques above: the potentially mysterious purpose of the interview.

5. Interactional roles and goals

In spite of the contrived purpose of linguistic interviews, the exchanges therein can be genuine. Interviewers and consultants often connect for a variety of social reasons and can occupy different roles within the context of the interview. The interviewer often takes the lead, eliciting social and linguistic information as well as sufficient speech for linguistic analyses. Participants may cooperate with interviewers by answering questions and speaking about appropriate topics, but they may also take on an agentive role within the interview and the research process. Interviews can be an opportunity to share information about their community, to convey beliefs, to reveal information or to discuss interests. Participants may initiate topics in line with their desire to share personal stories, opinions and cultural perspectives. Nevertheless, it is important to attend to the interactional aspects of the interview that can influence both the information conveyed by interviewees and the linguistic forms they use to do so.

Appreciating the goals of both the interviewer and the participant(s) is important in the interview process, particularly when eliciting qualitative data about cultural practice(s). Interviewees create and reinforce identities and positions in interviews and may have broader social and/or interactional motivations for presenting an attitude within the context of an interview. Wortham et al. (2011: 48–49) discuss ways in which interviewees affiliate with (and/or distance themselves from) opinions expressed in interviews. They may align with an interviewer's (perceived) view and enlist the interviewer as an ally or choose to project a potentially contrary position. Furthermore, interview questions that do not match their perspective of their cultural reality may rightfully annoy or frustrate participants. These issues underscore the importance of awareness of and continual reflection on the context of the community and the individual consultants.

6. Fieldworker profiles and approach

In previous sections, I have raised the issues related to the social backgrounds of fieldworkers; we turn now in this section to an explicit discussion of these issues. Social differences between interviewers and participants can contribute to complex dynamics between researchers and consultants. Interrelated social dimensions, such as socioeconomic status, gender, life stage, group membership and cultural background can influence interviewing, and fieldworkers must be aware of how they present themselves. There are many ways in which this can manifest; I highlight but a few.

Fieldworkers may have had more access to formal education than the speakers they interview, which can cause initial awkwardness in an interview despite the value of participants' expertise in their own language and culture. Aspects of identity, such as an age gap and/or differences in gender between fieldworkers and participants can also contribute to interactional aspects of the interview and to the types of variation used by speakers. Older speakers interviewed by younger fieldworkers may act as teachers or mentors, recalling experiences with the goal of enlightening or instructing. Differences in gender can also influence content and forms. For example, in my own recent work (Hoffman 2017), which is in turn part of a larger project on the English of different ethnolinguistic groups in Toronto (Hoffman and Walker 2010), young Italian-descent

men introduced narratives conveying their own toughness when interviewed by a female fieldworker from the same community, and used more retracted variants of /o/ (which is associated with a tough, masculine Italian identity [Boberg 2014]). This example also highlights the potential consequences of community membership on interviewing, since social distance and insider status can both affect interviews.

Both in-group and out-group fieldworkers can be successful interviewers, albeit with respective advantages (and disadvantages). Interviewers who share some social characteristics with a community may find their background advantageous. Many projects make use of community members for interviewing in language contact scenarios (Otheguy and Zentella 2012; Sharma 2011), and in research on code-mixing, community member interviewers may be best suited to the task. For example, the sociolinguistic competence of Torres Cacoullos and Travis's (2018) community fieldworkers was instrumental in eliciting their corpus of spontaneous Spanish-English code-switching in Texas. In-group fieldworkers may also have better access to and understanding of their communities, allowing for the elicitation of features marking ethnic and/or linguistic identity (Hoffman and Walker 2010). However, community membership can at times also disadvantage fieldworkers. It may be awkward for friends or family members to conduct interviews because of speakers' reluctance to answer questions about linguistic and cultural practices with which interviewers are obviously familiar. Consultants may assume that in-group fieldworkers are aware of local norms, and this may lead to misunderstandings and awkwardness (Mendoza-Denton 2008: 113−114).

Outsiders can be effective interviewers as well. I conducted the fieldwork in my research on the Spanish of Salvadorian youth in Toronto, Canada (Hoffman 2010). As a native English speaker of European descent, I had limited preexisting connections with Latina/o youth. I spent time in and around a community resource center getting to know some of the youth and their outreach coordinator. All in all, my fieldwork with the Salvadorian-Canadian youth was effective, despite my outsider status. My role allowed me to ask speakers about their community, situating them as cultural experts and tempering the awkwardness. Most of my consultants engaged me in narratives of their experiences, constructing their identities as Salvadorian and Latina/o, through both the content of our conversation and linguistic variation, though of course, my identity affected their language and what they told me.

Numerous other investigations of language contact have conducted successful fieldwork with interviewers who were outsiders (Cheshire et al. [2011] on Multicultural London English; Klee, Tight, and Caravedo [2011] on Spanish-Quechua contact; Nortier and Dorleijn [2008] on Moroccan Dutch). In this work, the interviewers were familiar with the communities and their cultural contexts, and had forged prior social connections with the groups. On the other hand, attitudes toward local varieties are also relevant to fieldwork in contact situations: speakers may be reluctant to use a variety with outsiders if they think it may be perceived as invalid. Some of my consultants distanced themselves from features they associated with Salvadorian Spanish, both in metalinguistic commentary and through variation (Hoffman 2010). Blainey (2017) notes that an interviewer's use of standard French intimidated speakers of Louisiana French (an obsolescing variety) and made them more reticent to participate in projects.

Outsider researchers interviewing jointly with insiders can not only benefit from a broader range of speech styles and social information but also draw on local expertise to enhance research design (Holmquist and Kahn 2017). Blainey (2017) highlights the

value of enlisting community members as coinvestigators to familiarize outsider researchers with community norms and to facilitate relationships. There are many ways to connect with communities: from contacting established social groups, local academic colleagues and/or community leaders to (gently) approaching seemingly gregarious individuals in public locations like cafés. Social media offer unprecedented access to potential participants. Connecting this way can be quicker and more efficient than in person; word can spread almost instantaneously within and between social networks. While recruiting will likely still include an in-person dimension to explain the work, social media and digital platforms can expedite initial contact. It is important to be mindful of the effect that these differing points of entry have on the perceptions of consultants.

7. Conclusions

Languages in contact offer dynamic contexts for the investigation of linguistic structures and the diverse, intricate ways speakers use language to convey social meaning. Interviewing is a fundamental tool for both quantitative and qualitative approaches to such research. This chapter has highlighted some approaches to interviewing as well as some issues to consider in language contact research. Ultimately, interviewers' openness to, sensitivity to and respect for consultants and their communities will support their success in this process.

8. References

Becker, Kara
2018 The sociolinguistic interview. In Christine Mallinson, Becky Childs & Gerard Van Herk (eds.), *Data collection in sociolinguistics*, 2nd edn., 99–107. New York, NY: Routledge.

Bell, Allan
1984 Language style as audience design. *Language in Society* 13. 145–204.

Bell, Allan & Gary Johnson
1997 Towards a sociolinguistics of style. *University of Pennsylvania Working Papers in Linguistics* 4(1). 1–21.

Besnier, Niko & Susan U. Philips
2014 Ethnographic methods for language and gender research. In Susan Ehrlich, Miriam Meyerhoff & Janet Holmes (eds.), *The handbook of language, gender, and sexuality*, 2nd edn., 123–140. Chichester, UK: Wiley Blackwell.

Blainey, Darcie
2017 Sociolinguistic research with endangered varieties: The case of Louisiana French. *Canadian Journal of Linguistics* 62(4). 576–595.

Boberg, Charles
2014 Ethnic divergence in Montreal English. *The Canadian Journal of Linguistics / La Revue Canadienne De Linguistique* 59(1). 55–82.

Briggs, Charles
1986 *Learning how to ask: A sociolinguistic appraisal of the role of the interview in social science research*. Cambridge: Cambridge University Press.

Chambers, J. K.
2008 Louis Gauchat, proto-variationist. *Historiographia Linguistica* 35(1/2). 213–225.

Cheshire, Jenny, Paul Kerswill, Sue Fox & Eivind Torgersen
2011 Contact, the feature pool and the speech community: The emergence of multicultural London English. *Journal of Sociolinguistics* 15(2). 151–196.
García Mouton, Pilar (coord.), Inés Fernández-Ordóñez, David Heap, Maria Pilar Perea, João Saramago & Xulio Sousa
2016 *ALPI-CSIC.* http://www.alpi.csic.es. (accessed on 10 September 2018). [Digital edition of Navarro Tomás, Tomás (dir.). 1933–1962. *Atlas Lingüístico de la Península Ibérica.* Madrid: CSIC.]
Gilliéron, Jules & Edmond Edmont
1969 *Atlas linguistique de la France*. Bologna: Forni.
Guardado, Martin
2002 Loss and maintenance of first language skills: Case studies of Hispanic families in Vancouver. *The Canadian Modern Language Review / La Revue Canadienne des Langues Vivantes* 58(3). 341–363.
Guardado, Martin & Ava Becker
2014 'Glued to the family': The role of familism in heritage language development strategies. *Language, Culture and Curriculum* 27(2). 163–181.
Gumperz, John
1964 Linguistic and social interaction in two communities. *American Anthropologist* 66(6). 137–153.
Hoffman, Michol F.
2010 *Salvadorian Spanish in Toronto: Phonological variation among Salvadorian Youth in a multilectal, multilingual context*. Munich: Lincom-Europa.
Hoffman, Michol F.
2017 In the front and in the back: The role of ethnicity in back vowel fronting in Toronto English. Paper presented at the Meeting of the American Dialect Society, Austin, Texas, 5–8 January.
Hoffman, Michol F. & James A. Walker
2010 Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change* 22. 37–67.
Holmquist, Jonathan & Hana Muzika Kahn
2017 Spanish and Kaqchikel-Maya: A study in town and village in Guatemala's central highlands. *International Journal of the Sociology of Language* 248. 3–24.
Hymes, Dell H.
1974 *Foundations in sociolinguistics: An ethnographic approach*. Philadelphia: University of Pennsylvania Press.
Jaberg, Karl, Jakob Jud & Paul Scheuermeier
1928 *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen: Ringier.
Johnstone, Barbara
2000 *Qualitative methods in sociolinguistics*. New York: Oxford University Press.
Klee, Carol A., Daniel G. Tight & Rocío Caravedo
2011 Variation and change in Peruvian Spanish word order: Language contact and dialect contact in Lima. *Southwest Journal of Linguistics* 30(2). 5–31.
Labov, William
1963 The social motivation of a sound change. *Word* 19. 273–309.
Labov, William
1966 *The social stratification of English in New York City.* Washington D.C.: Center for Applied Linguistics.
Labov, William
1972 *Language in the inner city: Studies in the Black English vernacular.* Philadelphia: University of Pennsylvania Press.

Labov, William
1984 Field methods of the project on linguistic change and variation. In John Baugh & Joel Sherzer (eds.), *Language in use*, 28–53. Englewood Cliffs, New Jersey: Prentice Hall.
Mendoza-Denton, Norma
2008 *Homegirls*: *Language and cultural practice among Latina youth gangs.* Malden, MA: Blackwell.
Mesthrie, Rajend
2012 Ethnicity, substrate and place: The dynamics of Coloured and Indian English in five South African cities in relation to the variable (t). *Language Variation and Change* 24(3). 371–395.
Meyerhoff, Miriam
2016 Methods, innovations and extensions: Reflections on half a century of methodology in social dialectology. *Journal of Sociolinguistics* 20(4). 431–452.
Navarro Tomás, Tomás (dir.)
1933–1962 *Atlas Lingüístico de la Península Ibérica.* Madrid: CSIC.
Nortier, Jacomine & Margreet Dorleijn
2008 A Moroccan accent in Dutch: A sociocultural style restricted to the Moroccan community? *International Journal of Bilingualism* 12(1). 125–142.
Otheguy, Ricardo & Ana Celia Zentella
2012 *Spanish in New York: Language contact, dialectal leveling, and structural continuity.* New York: Oxford University Press.
Platt, Jennifer
2012 The history of the interview. In Jaber F. Gubrium, James A. Holstein, Amir B. Marvasti & Karyn D. McKinney (eds.), *The Sage handbook of interview research: The complexity of the craft*, 2nd edn., 9–26. Thousand Oaks, CA: Sage.
Podesva, Robert
2007 Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics* 11. 478–504.
Poplack, Shanna
2017 *Borrowing: Loanwords in the speech community and in the grammar*. Oxford: Oxford University Press.
Sabino, Robin
2012 *Language contact in the Danish West Indies: Giving Jack his jacket*. Leiden: Brill.
Schilling, Natalie
2013 *Sociolinguistic fieldwork.* New York: Cambridge University Press.
Sharma, Devyani
2011 Style repertoire and social change in British Asian English. *Journal of Sociolinguistics* 15(4). 464–492.
Tagliamonte, Sali A.
2006 *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.
Tagliamonte, Sali A.
2016 *Making waves: The story of variationist sociolinguistics*. Malden & Oxford: Wiley-Blackwell.
Torres Cacoullos, Rena & Catherine E. Travis
2016 Two languages, one effect: Structural priming in spontaneous code-switching. *Bilingualism: Language and Cognition* 19. 733–753.
Torres Cacoullos, Rena & Catherine E. Travis
2018 *New Mexico Spanish-English Bilingual (NMSEB) corpus.* National Science Foundation 1019112/1019122. http://nmcode-switching.la.psu.edu. (accessed on 10 September 2018).
Valdés Kroff, Jorge R. & Matías Fernández-Duque
2017 Experimentally inducing Spanish-English codeswitching: A new conversation paradigm. In Kate Bellamy, Michael W. Child, Paz González, Antje Muntendam & M. Carmen

Parafita Couto (eds.), *Multidisciplinary approaches to bilingualism in the Hispanic and Lusophone world*, 211–234. Amsterdam & Philadelphia: John Benjamins.

Waltermire, Mark
2014 Language use and attitudes as stimuli for phonological change in border Uruguayan Spanish. In Dominic Watt & Carmen Llamas (eds.), *Language, borders and identity*, 70–89. Edinburgh: Edinburgh University Press.

Wolfson, Nessa
1976 Speech events and natural speech: Some implications for sociolinguistic methodology. *Language in Society* 5(2). 189–209.

Wortham, Stanton, Katherine Mortimer, Kathy Lee, Elaine Allard & Kimberly D. White
2011 Interviews as interactional data. *Language in Society* 40. 39–50.

Michol F. Hoffman, Toronto (Canada)

50. Participant Observation

1. Introduction
2. Participant observation: past and present
3. Observers observed: between surveillance and looking-talking back
4. Challenges: Multi-sited ethnography, auto-ethnography, collaboration, and performativity
5. Conclusion
6. References

1. Introduction

Participant observation is a method for collecting ethnographic (including linguistic) data. The basic idea is to embed oneself with the people one wishes to study, and to participate in their activities, so that one can begin to see the world through their eyes, and understand it from their point of view, while at the same time maintaining a critical, intellectual distance. As a method, it has been used especially by sociolinguists who were trained as linguistic anthropologists in the United States (e.g. Gal 1979; Meek 2010), but also by those who identify as third-wave variationists (e.g. Eckert 1989), and increasingly, by linguists engaging in language documentation (Dobrin and Schwartz 2016; for a discussion of the links between linguistics and ethnography, see Rampton, Maybin, and Roberts 2015). The advantage of participant observation (and ethnographic fieldwork more generally) is the ability to render 'thick descriptions' (Geertz 1973), that is, to provide accounts of social and linguistic life that foreground the lived experiences of individuals. Understanding sociocultural practices from the perspective of the insider is called the *emic perspective*. It is distinguished from the *etic perspective*, that is, the perspective of the outsider. Concepts such as 'language' or 'dialect', as well as speech acts such as 'greeting' or 'complimenting', thus cannot be universally defined; they are enacted and experienced in locally specific ways, and are imbued with different meanings in different sociocultural contexts.

https://doi.org/10.1515/9783110435351-050

This chapter does not provide a hands-on tutorial of how to 'do' participant observation; for this see the practical 'guides' published by, for example, Spradley (1980, reissued in 2016), DeWalt and DeWalt (2011), and Pelto (2016). Rather, I will provide a *critical* overview of participant observation, discussing its historical development and current status, as well as new approaches, such as autoethnography and performative fieldwork. The focus is on the method itself, rather than on its specific applications (e.g. to the study of language, social practices, and/or institutions). Participant observation was first developed in the field of anthropology, a discipline whose history is closely entangled with colonialism (Asad 1973; Gough 1968; Lévi-Strauss 1966; for a recent discussion, see also Nyamnjoh 2012). Critical reflexivity in anthropology has, since the late 1960s, interrogated the practice (and history) of 'doing fieldwork', and positioned it as an experience that is both individually and institutionally filtered. Doing participant observation thus comes with its own history, as well as with theoretical and epistemological challenges.

2. Participant observation: past and present

The term *participant observation* gained currency in the 1920s and 1930s, but the practice has a longer history. Regularly named in historical overviews as the first scholar to engage in such research is Frank Hamilton Cushing (1857-−1900), who studied the Zuni Indians in the 1880s. Cushing lived with the Zuni, observing and documenting their language and cultural practices. He was later initiated into the Priesthood of the Bow and given the name Tenatsali ('medicine flower'). Cushing called his approach the 'internal method': "a combination of linguistics, daily observation and intuition" (Hinsley 1983: 60). However, Cushing's impact on the discipline of anthropology was limited: his poetic and imaginative style, as well as his unwillingness to maintain critical distance, did not fit with the emerging professionalization of anthropology and its attempt to establish itself as a science. In many ways, Cushing was simply not 'disciplined' enough to allow for the discipline to recognize him.

It is Bronisław Malinowski (1884−1942) who is traditionally credited with having developed participant observation as a distinctive 'scientific', as well as 'professional', method of inquiry. It is a method that moves beyond interviews, instead foregrounding observation and listening, with the goal of understanding tacit and embodied (rather than explicit, discursively articulated) sociocultural knowledge. Core features of Malinowski's approach to fieldwork are: long-term participation in daily activities; learning the local language, dialect, or language variety; and engaging in everyday conversations as well as informal observations. Malinowski describes his 'method' (as he calls it) in *Argonouts of the West Pacific* as follows:

> Imagine yourself suddenly set down surrounded by all your gear, alone on a tropical beach, close to a native village, while the launch or dinghy which has brought you sails away out of sight […] Living in the village with no other business but to follow native life, one sees the customs, ceremonies and transactions over and over again, one has examples of their beliefs as they are actually lived through […] in this type of work, it is good for the Ethnographer sometimes to put aside camera, note book and pencil, and to join in himself in what is going on […] the final goal […] is, briefly, to grasp the native's point of view, his relation

> to life, to realise *his* vision of *his* world. (Malinowski ([1922] 2002: 3–18; emphasis in the original)

Fieldwork, according to Malinowski (2002: 16–18), produces three different types of texts: (i) analyses of social structures (based on "statistical documentation of concrete evidence"); (b) fine-grained, vivid and often literary descriptions of "the inponderability of actual life and typical behaviour"; and (iii) a *corpus inscriptionem* which documents "narratives", "typical utterances", and "items of folk-lore". While (i) is based on survey-type fieldwork, (ii) and (iii) are the outcome of participant observation.

Malinowski conducted his fieldwork on the Mailu and Trobriand Islands in a particular sociohistorical context: colonial rule and exploitation. Europeans have typically adopted a stance of 'discovery' towards the colonies: they were seen as far-away-and-strange lands, and research was linked to 'discovery' and 'adventure' (as evident in the subtitle of Malinowski [2002]). The trope of the-anthropologist-as-hero, who survives difficult conditions in the service of science and returns to 'tell the tale', was pervasive in these early accounts (Sontag 1963; Thompson 1995). Under colonial conditions, immersion was never more than partial: spending time among those who were consistently referred to as 'natives' (as in the quote above), or even as 'savages' (as in Malinowski [1929]), was not done under conditions of equality, and racial hierarchies shaped the fieldwork encounter throughout. Photos that document Malinowski's fieldwork show him as a man who, despite living among the Islanders and having learned Trobriand, continued to wear European clothing, and consequently created a strong, and highly visible, semiotic boundary between 'researcher' and 'researched'. A century later, Stebbins (2012: 303) reflects on the continued visibility of the white body in her fieldwork in Papua New Guinea: "I was always and already a *misis* (Tok Pisin for 'white woman', with all the colonialist implications the corresponding term *masta*, 'white man', more obviously denotes)". Thus, the observer effect (which has long troubled linguists and other social scientists) remains: researchers might be participating, but they remain visible, often hypervisible.

Malinowski's legacy became highly controversial when his diary was published in 1967: the text is peppered with racial epithets, and Malinowski expresses ambivalence, bitterness, and also disrespect towards the people he worked with. Geertz (1983: 56) commented that the diary demolished "the myth of the chameleon fieldworker, perfectly self-tuned to his exotic surroundings, a walking miracle of empathy, tact, patience and cosmopolitanism" (see also Lamont 2014). While the publication of the diary cast a lasting shadow over the legacy of Malinowski, it did not dampen the importance which anthropology attached to fieldwork in general, and to participant observation in particular: it remains a methodology of choice and a professional rite-of-passage for emerging scholars in the discipline (Agar 1980). It is worth noting that early linguistic anthropologists, such as Franz Boas (1858–1942) and Edward Sapir (1883–1939), did not engage in participant observation: both collected data primarily through elicitation from a few key people (Murray 1994; Wax 1971). Participant observation entered linguistics primarily through the work of Gumperz and Hymes in the 1970s (see for example Gumperz and Hymes 1972; Hymes 1978).

The colonial history of participant observation remains relevant to contemporary discussions. This became visible in the controversies surrounding Goffman's (2014) urban ethnography about the hyper-policing of an inner-city neighborhood (which she calls

6th Street). Her book became, in the words of the journalist Lewis-Kraus (2016), "a proxy for old and unsettled arguments about ethnography that extended far beyond her own particular case". Methodologically, Goffman (2014: 240), like many other urban sociologists, adopted participant observation: "I wanted to be a participant observer. I wanted to walk alongside Mike and his friends and neighbors so that I could understand their everyday worries and small triumphs from the inside". The book was initially well-received, and even described as a classic-in-the-making. But soon critical voices emerged. Inconsistencies in the text cast doubt on the accuracy of the claims made, and the very method of participant observation came under critique. The intensity of her participation, which involved the willingness to carry out criminal acts, was considered unethical by some, who emphasized the need to maintain critical distance (on violence and other challenges of fieldwork, see Nordstrom and Robben [1995], as well as Kulick and Willson [1995]). Participant observation, in other words, is always in a double-bind, and wishing to observe as well as to participate is a delicate balancing act: Malinowski erred on the side of observation, Cushing and (more than a century later) Goffman on the side of participation. Others were concerned about the racial politics and power dynamics that informed her work, that is, the ways in which the text was seen to sensationalize black poverty, violence, and crime (Rios 2015). Betts (2014) takes Goffman to task for perpetuating the same tropes that informed early anthropology, writing that "[w]hat offends me is that Goffman has turned 6th Street into a jungle that she has braved", and Sharpe (2014) commented that the book is simply "another story about a white lady come to study young black men". This criticism of Goffman's work shows how the legacies of anthropology's colonial roots live on in the present. There is, firstly, the trope of the ethnographer-as-hero who survived dangerous encounters in the service of the discipline, and secondly, a focus on societies and peoples that are maximally 'other' from the mostly (and traditionally) white academic readership (that is, non-modern, non-European, non-bourgeois societies or peoples are the preferred object of study). In conducting work Malinowski-style, anthropologists inevitably encounter the 'problem of representation', that is, the question of who should write about whom (and for whom). Importantly, the problem of representation is not simply a moral question (an 'ought-to' question), but also concerns epistemology, or our ability to know. In other words, if the study of 'the other' is closely linked to histories of oppression and violence, then the question arises whether knowing is ever possible under such conditions (Maldonado-Torres 2007). While it is generally true that the portrayals produced by outsider researchers will always differ from those of insider researcher, this difference takes on a particular shape in contexts where the relationships between researcher and researched are of a highly hierarchical nature. It is in such contexts that we find instances of 'epistemic violence' (Spivak 1988). In other words, knowledges about 'the other' that are produced in contexts shaped by sociohistorically persistent inequalities are likely to contain stereotypes, distortions, and misrepresentations. Consequently, participant observation requires not simply a willingness to embed oneself into a different society, but also (and quite centrally) a high degree of reflexivity, considering carefully the researchers' own motivations, blindspots, and positionalities (i.e. their location in time-space, as well as their raced and gendered selves). This affects not only fieldwork and data collection, but also interpretation and writing (Clifford and Marcus 1986).

3. Observers observed: between surveillance and looking-talking back

Consider what happens if one changes the direction of the gaze: not looking at 'the other', 'the people' to be studied and researched, but looking at those who do the studying, at the researchers themselves. In Stocking's (1983) *Observers Observed*, anthropologists observe themselves through a critical reevaluation of the discipline's history. Another example of 'observers observed' is Storey's (2014) discussion of the Mass-Observation project, which was established in the United Kingdom in 1937. Storey's analysis focuses on the first phase of Mass-Observation (up until 1950), and not the second project (which started in the early 1980s). The Mass-Observation project allows us to think not only about the observers, but also about the very idea of observation, and how this practice can border on surveillance. That observation is closely linked to control and governmentality becomes clear when we consider censuses, surveys, and taxation records, which have served to observe and record people for centuries. In the 20th century, with cameras and tape recorders, new methods of augmenting observation emerged: not just notetaking, but also audio and video recordings, and photographs.

Mass-Observation was inspired by colonial anthropology, and indeed, the work of Malinowski. However, it moved beyond the lone ethnographer, and instead massified observation: in principle, everyone could be an observer. The only rule was to be invisible, to embed oneself in the situation, so that one was no longer noticed, and then to look, listen, and record. The aim of the project was to produce a popular 'anthropology of ourselves', to study that which is familiar (the everyday rather than the exotic). In this sense, the project anticipated field sites that are increasingly common in contemporary research: no longer far away islands, but rather neighborhoods, hospitals, airports, schools, and so forth (Gupta and Ferguson 1997). Yet despite the intention to massify observation, to turn it into an everyday practice − indeed, to democratize and de-professionalize it − a gulf emerged, yet again, between 'the observed' (mostly the British working classes) and 'the observers' (mainly members of the upper-class, often Oxbridge students). This was particularly visible in the study of everyday life in Bolton. Many of the observers, such as the photographer Spender, came from well-to-do backgrounds ("with nannies and governesses" [Storey 2014: 50]), while others belonged to the growing middle-class, including teachers, clerks, and commercial travelers. Their class privilege allowed them to engage with the project (by having sufficient leisure time), and inevitably shaped their interactions and reports. Thus, they produced an ethnography of working class lives through the eyes of middle class observers. It made the working classes "an object of middle-class knowledge" (Storey 2014: 49), just like early anthropology had made the lives of the racial 'other' an object of colonial knowledge.

However, neither the working classes, nor any other groups that have been studied by researchers, have remained quiet. In *Black Looks: Race and Representation,* hooks (1992: 165) notes that it is not only white people who study and observe black people, but also black people who collect "details, facts, observations, and psychoanalytic readings of the white Other". With respect to the situation in the United States, hooks (1992: 165) continues: "[a]lthough there has never been any official body of black people in the United States who have gathered as anthropologists and/or ethnographers to study whiteness, black folks have, from slavery on, shared in conversations with one another 'special' knowledge of whiteness gleaned from close scrutiny of white people".

As one conducts fieldwork, one is inevitably embedded in what bell hooks (1992) calls 'looking relations': the researcher looks, but the researched always look back. In her work, hooks asserts not only the power of the gaze, but also reminds us of the fundamentally relational and interactive nature of all fieldwork: as ethnographers are interacting and trying to make sense of what they observe, those being observed are also interacting and trying to make sense of the observers. In a seminal volume on the ethics of linguistic fieldwork, Cameron et al. (1992: 5) move in a similar direction, emphasizing that "research subjects […] are active and reflexive beings who have insights into their situations and experiences", including the situation of 'being researched' (see also Jaspers and Meeuwis [2013] on relationality in fieldwork). Resistance to being researched, to being the object of the expert's gaze, is nothing new. In his introduction to an edited collection of Malinowski's work on the Trobriand Islands, Young (1979: 15) draws on the reminiscences of Father Baldwin, a missionary in the area, who was a keen amateur anthropologist. Baldwin suggests that while Malinowski certainly grasped something ("facts"), he failed to capture the actual richness of experience, leaving the people he studied angry and resentful. According to Young (1979: 15), Baldwin wrote: "I was surprised at the number of times informants helping me with checking Malinowski would bridle […] They did not quarrel with facts or explanation, only with the colouring as it were. The sense expressed was not the sense they had of themselves".

Fifty years after the publication of Malinowski's first monograph, Kasaipwalova, a poet and political activist from the Trobriand Islands, echoes this resentment at being misrepresented. His reflections (as quoted in Young 1979: 17; emphasis in the original) point to the lingering presence of colonial anthropology in postcolonial discourses:

> I have felt many times very embarrassed and awkward when I meet new people and they ask me, 'where are you from?', and I say, 'Oh, I am from the Trobriand Islands', and they reply excitedly, 'Oh, Malinowski free love!!' What I am pointing out is that if we are going to depend on anthropological studies to define our history and our culture and our 'future', then we are *lost*.

Discourses of resistance to being studied became especially visible in the 1990s. In disability studies, for example, the slogan 'nothing about us without us' rallied against outsider research of disability issues, and an antipathy toward academic researchers came to be openly expressed in various contexts. Harvey (2004: 172), an ethnographer of religion, attending a traditional social gathering of the Mi'kmaq Nation in Newfoundland, remembers a tongue-in-cheek, yet serious, interaction which articulates this antipathy: "At lunch […] I asked people surrounding a vacant chair if I could join them. They said, 'Are you an anthro? […] because we eat anthros'". A seminal, book-length discussion of 'talking/writing/researching back' is Smith's ([1999] 2012) *Decolonizing Methodologies*. Like hooks, Smith (2012: 1) takes note of the many stories "about observers" that circulate in indigenous communities, and how the very word "research" when mentioned "stirs up silence, it conjures up bad memories, it raises a smile that is knowing and distrustful" (see also Nandy 1983; Wilson 1992). Attempts to include voices other than the researcher's bring with them their own moral dilemmas, rooted in the complex and unequal geopolitics of knowledge that go back to colonialism, but persist into the present day. This was illustrated poignantly in a coauthored paper by McLaughlin, the American expert-linguist, and Sall, the so-called 'native-speaker assistant' in Senegal (2001). At the beginning of their paper, which reflects on linguistic fieldwork

from their different perspectives, stands a short dialogue, which forcefully reminds us that power relations persist even in what we might call 'dialogic ethnography' (Clifford 1983).

> *Sall:* Somehow one has the impression that we are always the object and never the subject. We are the "material" that toubabs [white people] come to study.
> *McLaughlin:* But this time, by presenting your own narrative, don't you think that you have the opportunity to be the subject instead of the object?
> *Sall:* (Laughter) I talk about myself, but only at your initiative. Where does that put us? (McLaughlin and Sall 2001: 189)

4. Challenges: multi-sited ethnography, autoethnography, collaboration, and performativity

The question remains how we can articulate a version of participant observation that moves beyond the colonial origins of the discipline, and whether we need different methods for a different anthropology. Questions such as these have come to the fore since the 1960s, and various challenges to participant observation have been formulated.

Growing interest in globalization and transnationalism since the late 20th century has led anthropologists to reexamine the idea that long-term, intense fieldwork should be carried out within specific, placed communities. In order to account for mobility, the fragmented nature of contemporary societies, and the fact that local practices are shaped by global influences, multi-sited ethnography emerged in the 1990s (Marcus 1995). In multi-sited ethnography, the researcher follows people, objects, networks, and texts across space, and fieldwork tends to be conducted for shorter periods of time and in spatially noncontiguous locations. In such work, it is difficult to maintain the insider-outsider distinction: the researcher may be an insider in some spaces, but an outsider in others.

A desire to explore the complexity of researcher identities more fully − and beyond the insider-outsider dichotomy − has influenced the development of another approach: autoethnography. Taking seriously the argument that positionality affects the ways in which one conducts research, autoethnography incorporates the researcher's/writer's autobiographical data into the analysis and interpretation of a given topic (Chang 2008; Ellis 2004; Jones, Adams, and Ellis 2013). From this perspective, research is not simply about the other seeing the world from their point of view, but also about the self: understanding happens in the complex interactional space between self and other. Autoethnography, just like multi-sited ethnography, disrupts the insider-outsider dichotomy: I might be an outsider to the people I study, but by taking my own experiences, limitations, and thoughts into account and making them part of the texts that I produce, I am also an insider to my own experience.

To turn the gaze away from the other and towards the self, to engage in a participant-observation of the observer, as well as of the power relations, exploitations, and entanglements that define the ethnographic encounter, is core to Radebe's (2016) reflections on 'decolonizing anthropology'. In addition, decolonial approaches have emphasized the need for more collaborative, and indeed collective, methodologies in research which disrupt and dismantle the roles of researcher-researched and observer-informant. In such

projects, there is a sense of 'walking together' and 'writing together' as the work takes shape, is carried out, and is prepared for publication (Rappaport 2008; Solano and Rappaport 2011). Proposing the need for the "complete transformation of the anthropologist's craft", Vasco Uribe (2011: 19) argues that it is not enough to revise theories, but that changing methodologies is central to the transformation of the discipline. For Vasco Uribe (2011: 29), research is a type of dialogue, which can almost be a confrontation: a confrontation between academic and popular ways of knowing, a confrontation in which both types of knowledge are "validated, refined, and combined, to produce concepts, methods and procedures for activist research (*investigación-acción*)". Activism, the transformation of an unequal reality, is core to decolonial methodologies, which are positioned within a larger project of social justice. An activist-collective approach challenges the idea of individual researchers as 'translators' of facts that were reported to them by numerous 'others'. Knowledge is instead located in interaction: it is emergent rather than given. Collaboration, in this sense, is knowledge-coproduction. Vasco Uribe (2011: 32) suggests that the authority and control of the anthropologist can only be broken, if fieldwork is based on the following principles: "local participants must assert control over the research […] the ethnographer's opinion is but one of many […] the ethnographer's opinion must be shared with local participants […] local participants must formulate their own research proposals […] oral narratives must be accepted as truths, and not as mere discourses". Yet Sall's admonition (cited above) still stands: "I talk about myself, but only at your initiative. Where does that put us?" (McLaughlin and Sall 2001: 189). This encourages us, from the very beginning of any project, to reflect on who gives voice to whom, and when.

A different take on the insider-outsider dichotomy is offered by work on performance and performativity in fieldwork. Castañeda (2006) speaks about the 'invisible theatre of fieldwork'. He draws on Boal's (1992) 'theatre of the oppressed', and in particular his idea of the 'Spec-Actors': that is, the way in which we are all simultaneously actors and observers in everyday life, triggering actions and reactions as we move between these two roles. We are all participant-observers! Because participant observation appropriates the techniques of the everyday (talking, looking, interacting, asking), it automatically takes part in the performance and performativity of the everyday (and this in turn is what makes it so difficult for researchers to know when they are 'doing' fieldwork, and when they are 'not doing' fieldwork). While the theatre of fieldwork, especially in the form of participant observation, is 'invisible' in the sense that it is not staged as 'theatrical', it nevertheless raises important questions about performance and stage: Who performs for whom? Who observes whom? For Castañeda, the prime audience of ethnographic work is not the reader of the academic book or article, but the diverse and constantly shifting audience(s) that emerge in the fieldwork encounter. The performative dimensions of fieldwork are also theorized by Storch (2018) in her ongoing work on mass tourism. Drawing on the sociological idea of 'role', she too conceptualizes fieldwork in (invisible) theatrical terms (akin to Goffman's [1959] work on *The Presentation of Self in Everyday Life*). What is the role of the researcher in the theater of the everyday? And what would happen if one accepts the role that is given to us by others? Thus, when studying mass tourism sites, one is rarely visible as an academic; rather, and depending quite clearly what one looks like, one might be positioned as a tourist, customer, or guest. Sometimes one remains unseen, like an unimportant background figure, sometimes one is perceived as attractive or annoying, and so forth. To acknowledge the complex

interactional socialities in which fieldwork takes place means that we relinquish − consciously and deliberately − control over 'our fieldwork', and that we accept its inevitable messiness, opening ourselves up to different forms of relations (e.g. dysfunctional, oppressive, nourishing, happy, colonial, or anticolonial relations). The textual equivalents to this are not cohesive and large-scale expert analyses, but what Storch calls 'small stories'. Like Vasco Uribe and Castañeda, Storch emphasizes how fieldwork is always an intervention, or a provocation, which shares a strong family resemblance with ethnomethodological breaching experiments (Garfinkel 1963). Maybe we should not think of the anthropologist as an 'expert' who comes to study a 'culture', 'community', 'language', or 'practice', but rather as a version of the trickster, the archetypical performer who plays with the social order (including language). This could open new ways for doing fieldwork, and perhaps inaugurate a version of what Castañeda calls 'experimental fieldwork'. In this way, ethnography can finally relinquish the old social science conundrum, the so-called 'observer paradox', the desire for invisibility and 'pure' data, and can accept that purity and authenticity are but red herrings, and that anything that happens, whether typical or atypical, whether anecdotally observed or systematically studied, matters in sociocultural life (on the importance of anecdote, see especially Gallop 2002; Simons 1978).

5. Conclusion

Participant observation has long defined anthropology as a discipline: anthropologists were university-trained specialists who developed theories of societies and peoples based on long-term immersive fieldwork. Participant observation was, in the words of Clifford (1983: 122), "a professional norm". However, since the 1960s, and ever more forcefully since the 1990s, participant observation has come under critique as colonial in origin, and as perpetuating unhelpful discourses of self and other. Increasingly, scholars have recognized the political and intersubjective 'messiness' of participant observation, and emphasized the need for critical reflexivity. Participant observation remains, at this point in time, a prominent method taught at universities, practiced by anthropologists, and topicalized in textbooks. It continues to be an attractive and important method, as it allows researchers to gain insight into community practices (linguistic and otherwise) and to reduce the observer's effect, allowing for the collection of rich data-in-context (Dobrin and Schwartz 2016). However, the method is not uncontroversial, and well-theorized alternatives exist for those who wish to move beyond participant observation: multi-sited ethnography, autoethnography, collaborative fieldwork, and performative fieldwork. The debate is open.

6. References

Agar, Michael
1980 *The professional stranger: An informal introduction to ethnography.* University of Michigan: Academic Press.

Asad, Talal (ed.)
1973 *Anthropology and the colonial encounter.* London: Ithaca Press.

Betts, Dwayne
2014, 10 July The stoop isn't the jungle. Slate. http://www.slate.com/articles/news_and_politics/jurisprudence/2014/07/alice_goffman_s_on_the_run_she_is_wrong_about_black_urban_life.html. (accessed on 16 August 2018).
Boal, Augusto
1992 *Games for actors and non-actors*. London: Routledge.
Cameron, Deborah, Elizabeth Frazer, Penelope Harvey, M. B. H. Rampton & Kay Richardson
1992 *Researching language: Issues of power and method*. London: Routledge.
Castañeda, Quetzil E.
2006 The invisible theatre of ethnography: Performative principles of fieldwork. *Anthropological Quarterly* 79. 75–104.
Chang, Heewong
2008 *Autoethnography as method*. London: Routledge.
Clifford, James
1983 On ethnographic authority. *Representations* 2. 118–146.
Clifford, James & George E. Marcus (eds.)
1986 *Writing culture: The poetics and politics of ethnography.* Berkeley: University of California Press.
DeWalt, Kathleen M. & Billy R. DeWalt
2011 *Participant observation: A guide for fieldworkers*, 2nd edn. Lanham: Altamira.
Dobrin, Lise M. & Saul Schwartz
2016 Collaboration or participant observation: Rethinking models of 'linguistic social work'. *Language Documentation and Conservation* 10. 253–277.
Eckert, Penelope
1989 *Jocks and burnouts: Social categories and identity in the high school*. New York/London: Columbia University, Teachers College.
Ellis, Carolyn
2004 *The ethnographic I: A methodological novel about autoethnography.* Lanham: Altamira.
Gal, Susan
1979 *Language shift: Social determinants of linguistic change in bilingual Austria*. University of California: Academic Press.
Gallop, Jane
2002 *Anecdotal theory.* Durham: Duke University Press.
Garfinkel, Harold
1963 A conception of and experiments with 'trust' as a condition of concerted stable actions. In O. J. Harvey (ed.), *Motivation and social interaction: Cognitive approaches*, 187–238. New York: Ronald Press Co.
Geertz, Clifford
1973 *The interpretation of cultures: Selected essays*. New York: Basic Books.
Geertz, Clifford
1983 *Local knowledge: Further essay in interpretive anthropology.* London: Basic Books.
Goffman, Alice
2014 *On the run: Fugitive life in an American city.* Chicago/London: University of Chicago Press.
Goffman, Erving
1959 *The presentation of self in everyday life*. London: Penguin.
Gough, Kathleen
1968 New proposals for anthropologists. *Current Anthropology* 9(5). 403–407.
Gumperz, John & Dell Hymes (eds.)
1972 *Directions in sociolinguistics: The ethnography of communication.* Oxford: Blackwell.

Gupta, Akhil & James Ferguson
1997 *Anthropological locations: Boundaries and grounds of a field science.* Berkeley: University of California Press.
Harvey, Graham
2004 Performing and constructing research as guesthood in the study of religions. In Lynne Hume & Jane Mulcock (eds.)*, Anthropologists in the field: Cases in participant observation*, 168–182. New York: Columbia University Press.
Hinsley, Curtis
1983 Ethnographic charisma and scientific routine: Cushing and Fewkes in the American South West, 1879–1893. In George W. Stocking (ed.), *Observers observed: Essays on ethnographic fieldwork*, 53–69. Madison: University of Wisconsin Press.
hooks, bell
1992 *Black looks: Race and representation*. Boston: Southend Press.
Hymes, Dell
1978 What is ethnography? *Sociolinguistic Working Papers* 45. Austin, Texas: Southwest Educational Development Laboratory, Indiana University.
Jaspers, Jürgen & Michael Meeuwis
2013 Away with linguists! Normativity, inequality, and metascientific reflexivity in sociolinguistic fieldwork. *Multilingua* 32(6). 725–749.
Jones, Stacey Holman, Tony E. Adams & Carolyn Ellis (eds.)
2013 *Handbook of autoethnography*. London: Routledge.
Kulick, Don & Margaret Willson (eds.)
1995 *Taboo: Sex, identity, and erotic subjectivity in anthropological fieldwork.* London: Routledge.
Lamont, Mark
2014 Malinowski and the 'Native Question'. In Regna Darnell & Frederic W. Gleach (eds.), *Anthropologists and their traditions across national borders*, 69–110. Lincoln & London: University of Nebraska Press.
Lévi-Strauss, Claude
1966 Anthropology: Its achievements and future. *Current Anthropology* 7(2). 124–127.
Lewis-Kraus, Gideon
2016, 12 January The trials of Alice Goffman. *The New York Times Magazine*. https://www.nytimes.com/2016/01/17/magazine/the-trials-of-alice-goffman.html. (accessed on 16 August 2018).
Maldonado-Torres, Nelson
2007 On the coloniality of being: Contributions to the development of a concept. *Cultural Studies* 21(2–3). 240–270.
Malinowski, Bronisław
1929 *The sexual life of savages in North-Western Melanesia: An ethnographic account of courtship, marriage, and family life among the natives on the Trobriand Islands, British New Guinea* (with a preface by Havelock Ellis). New York: Horace Liveright.
Malinowski, Bronisław
1967 *A diary in the strict sense of the term* (with an introduction by Raymond Firth). London: Routledge and Kegan Paul.
Malinowski, Bronisław
2002 [1922] *Argonouts of the West Pacific: An account of native enterprise and adventure in the archipelagoes of Melanesian New Guinea* (with a preface by Sir Simon Frazer). London: Routledge.
Marcus, George E.
1995 Ethnography in/of the world system: The emergence of multi-sited ethnography. *Annual Review of Anthropology* 24. 95–117.

McLaughlin, Fiona & Thierno Seydou Sall
2001 The give and take of fieldwork: Noun classes and other concerns. In Paul Newman & Martha Ratliff (eds.), *Linguistic fieldwork*, 189–210. Cambridge: Cambridge University Press.
Meek, Barbra A.
2010 *We are our language. An ethnography of language revitalization in a Northern Athabaskan community.* Arizona: University of Arizona Press.
Murray, Stephen O.
1994 *Theory groups and the study of language in North America: A social history.* Amsterdam: John Benjamins.
Nandy, Ashis
1983 *The intimate enemy: Loss and recovery of self under colonialism*. Oxford: Oxford University Press.
Nordstrom, Carolyn & Antonius C. G. M. Robben (eds.)
1995 *Fieldwork under fire: Contemporary studies of violence and survival*. Berkeley: University of California Press.
Nyamnjoh, Francis B.
2012 Potted plants in greenhouses: A critical reflection on the resilience of colonial education in Africa. *Journal of Asian and African Studies* 47(2). 129–154.
Pelto, Pertti J.
2016 *Applied ethnography: Guidelines for field research*. London: Routledge.
Radebe, Zodwa
2016, 23 May On decolonising anthropology. https://savageminds.org/2016/05/23/on-decolonising-anthropology/. (accessed on 16 August 2018).
Rampton, Ben, Janet Maybin & Celia Roberts
2015 Theory and method in linguistic ethnography. In Julia Snell, Sara Shaw & Fiona Copland (eds.), *Linguistic ethnography: Interdisciplinary approaches*, 14–50. Basingstoke & New York: Palgrave Macmillan.
Rappaport, Joanne
2008 Beyond participant observation: Collaborative ethnography as theoretical innovation. *Collaborative Anthropologies* 1. 1–31.
Rios, Victor M.
2015 *On the run: Fugitive life in an American city* by Alice Goffman. *American Journal of Sociology* 121(1). 306–308.
Sharpe, Christina
2014, 8 August Black life, annotated. *The New Inquiry.* https://thenewinquiry.com/black-life-annotated/ (accessed on 16 August 2018).
Simons, Herbert W.
1978 In praise of muddleheaded anecdotalism. *Western Journal of Communication* 42(1). 21–28.
Smith, Linda Tuhiwai
2012 [1999] *Decolonizing methodologies: Research and indigenous peoples*, 2nd edn. London & New York: Zed.
Solano, Xochitl Leyva & Joanne Rappaport
2011 Walking and doing: About decolonial practices. *Collaborative Anthropologies* 4. 119–138.
Sontag, Susan
1963 A hero of our time. *New York Review of Books* 1(7). 6–8.
Spivak, Gayatri Chakravorty
1988 Can the subaltern speak? In Cary Nelson & Lawrence Grossberg (eds.), *Marxism and the interpretation of culture*, 271–313. Chicago: University of Illinois Press.
Spradley, James P.
2016 [1980] *Participant observation*, 2nd edn. Long Grove, Illinois: Waveland Press.

Stebbins, Tonya N.
2012 On being a linguist and doing linguistics: Negotiating ideology through performativity. *Language Documentation and Conservation* 6. 292–317.
Stocking, George W. (ed.)
1983 *Observers observed: Essays on ethnographic fieldwork*. Madison: University of Wisconsin Press.
Storch, Anne
2018 Small stories. *The Mouth* 2. 97–117.
Storey, John
2014 *From popular culture to everyday life*. London: Routledge.
Thompson, Christina A.
1995 Anthropology's Conrad: Malinowski in the tropics and what he read. *The Journal of Pacific History* 30(1). 53–75.
Vasco Uribe, Luis Guillermo
2011 Rethinking fieldwork and ethnographic writing. *Collaborative Anthropologies* 4. 18–66.
Wax, Rosalie
1971 *Doing fieldwork: Warnings and advice.* Chicago & London: University of Chicago Press.
Wilson, Ken
1992 Thinking about the ethics of fieldwork. In Stephen Devereux & John Hoddinott (eds.), *Fieldwork in developing countries*, 179–199. Boulder: Lynne Rienner.
Young, Michael W. (ed.)
1979 *The ethnography of Malinowski: The Trobriand Islands 1915–1918*. London: Routledge and Kegan Paul.

Ana Deumert, Cape Town (South Africa)

51. Discourse analysis

1. Introduction
2. Approaching discourses
3. Foucauldian discourse analysis
4. Analyzing discourse
5. Taking stock with a perspective
6. References

1. Introduction

In 1988, Peter H. Nelde commented on future directions for 'contact linguistics'. As one of the central proponents of this research tradition, he argued that 'conflict' should be regarded as an inherent feature of contact between languages, and more importantly, between the speakers of these languages (see also Darquennes 2015). While listing various reasons and constellations that potentially lead to situations of conflict (e.g. related to linguistic minorities, bilingualism, or migration), he urged researchers to understand language conflict first and foremost as mirroring underlying sociopolitical asymmetries and inequalities of language groups that are in contact with each other (Nelde 1988: 79–80).

https://doi.org/10.1515/9783110435351-051

Conflicts, as well as related negotiations about inequalities between speakers, become apparent in discourses, and more importantly, as I argue in this chapter, in their analyses. In line with Nelde's critical perspective, I tease out the relevance of discourse analysis as a method for studying language contact, multilingual settings, and related conflicts between languages (or rather: speakers) in contact. Before turning to concrete examples of such analyses and studies in order to understand specific contact situations and their consequences for speakers, I discuss different epistemological traditions of discourse analysis. Finally, I point to potential shortcomings of discourse analysis in this context.

2. Approaching discourses

As is regularly noted, the term 'discourse' has differing usages and meanings. While in everyday talk and in most media usage, 'discourse' normally stands for the sum of statements on a certain topic, its meaning in scientific contexts is epistemologically more complex, depending on the ontological understanding of language. Used and applied differently by various scientific disciplines − and with varying approaches even within disciplines − it has become commonplace to comment on the terminological vagueness and conceptual indeterminacy of 'discourse'. However, despite the undeniable 'flexibility' of the term, it remains a central concept for analyzing the production, formation, and circulation of knowledge in society, while foregrounding ideologies and mechanisms of power. Yet it should be kept in mind that there is not one form of discourse analysis, but rather there are different ontologies, which lead directly to the existence of two main 'camps' in linguistic discourse analysis. These two camps have broader implications, as they are also the two ontologies defining the development of linguistics proper: formalism and functionalism. The formalist approach to discourse analysis in linguistics is mainly concerned with discovering and describing structures and/or categories that emerge in (predominantly written) interaction. In this instance, discourse is understood as an entity "above the sentence" (Schiffrin 1994: 23, 1996), thereby correlating with or complementing 'text'. In contrast, functionalism ascribes specific functions to language use, which is thus related *a priori* to social practice and social context. Under this functionalist paradigm, discourse is understood as related to situated language patterns that can take various forms or modalities (Blommaert 2005; Cameron 2001).

3. Foucauldian discourse analysis

From the space between these two ontologies, it could be argued that yet another ontology has emerged (Duchêne 2008; Spitzmüller and Warnke 2011b) which regards language as inherently and inextricably linked with social practices, and is inspired by Foucault's writing on discourses (Foucault 1969, 1971, 1981) and power (Foucault 1975). Researchers in this tradition have an analytical understanding of discourses as the socially stratifying collective knowledge of specific cultures or groups (Spitzmüller and Warnke 2011b: 8), and have an explicit interest in its materiality (Duchêne 2008: 30). Such collective knowledge, it is argued, emerges in systematic and systemic formations of meaning, and materializes in diverse yet coherent forms of expressions, practices, and realities. As

such, discourses are inherently related to the production of 'knowledge' and regimes of 'truth'. In situating 'knowledge' and 'truth' as historical products of meaning formation, discourse analysis that applies 'critique' (in a Foucauldian sense) aims to debunk claims of natural, universal, and/or timeless validity, and thereby to unveil the power structures that determine access to certain positions and resources in society.

On the one hand, this means that 'reality' is not only socially, but also discursively, constructed; on the other, critical approaches consider discourses not only as constituting and constitutive of social practices, but as inextricably infused with power structures, e.g. in *Critical Discourse Analysis* (or CDA, e.g. Wodak and Meyer [2009a, 2009b]; and related, *Discourse-Historical Approach*, or DHA, e.g. Dorostkar [2014]; Reisigl and Wodak [2009]) or *critical sociolinguistics* (Duchêne 2008; Heller 2002). 'Critique' can also be understood in a more applied way, as e.g. by researchers positioning themselves in the tradition of Critical Discourse Analysis (CDA). For them, taking a critical stance essentially means analyzing power relations that they presume to be asymmetrical, as well as (self-)reflecting on the researcher's position, ultimately leading to emancipation (cf. Wodak and Meyer 2009a: 6–7, 2009b). While Foucauldian discourse analysts thus shy away from claims to 'truth', the uncovering and proposition of such 'truths' in the form of output (e.g. recommendations, training workshops, and/or teaching materials) are part of the CDA agenda (Reisigl and Wodak 2009).

Now, if discourses are interconnected with power and truth, the possibility of contributing to a discourse is far from arbitrary; there is a logic as to why certain contributions are treated as valuable, reasonable, and/or legitimate, while others are treated as nonsensical, irrational, and/or illegitimate (Martín Rojo 2017). In other words, a discourse disposes of rules and structures that determine the production of utterances and texts in order to create an internal coherence, which is why Foucault (1971) invokes the idea of a 'discursive order'. For him, both function and form are of relevance: not only *what* is said is of interest, but also *how* it is being said and by *whom*. Yet in spite of their regulated internal coherence, discourses are far from monolithic or static. They are prone to changes, ruptures, and shifts. Those actors and institutions whose agendas and interests are reflected in these changes and shifts are the ones considered to be powerful in a Foucauldian sense (Duchêne 2008; Heller 2002), while the ones whose voices are left unconsidered (or are even considered irrational) are the ones without power (Blommaert 2005; Busch 2016).

In short, an analysis of the discursive order and related power structures is of particular relevance in situations of contact, as these inherently bear the potential for conflict and are imbued with specific structures of power and distributions of resources. These distributions are often unequitable, which in turn has material consequences for speakers of the different languages or varieties in contact.

4. Analyzing discourse

The question remains how to analyze discourse. The following section presents examples of empirical applications which might make use of different forms of data, but which are based on a Foucauldian understanding of discourse (i.e. of an internal coherence of and within utterances and texts, as well of a larger historical framework in which utterances and texts are embedded). In the end, these examples should tease out the analytical advantages of discourse analysis for contact linguistics.

4.1. A discursive analysis of interviews

Interviews are a common research method to access information, opinions, or ultimately, discourses. Just as any other text or practice, interviews are embedded in historical contexts with a particular discursive order. What people will say – and how – in an interview, and how they position themselves, is thus indexical of the internal coherence and borders of a discourse. For an empirical illustration of how to approach discourses analytically through interviews, I discuss the example of 'interpretative repertoires' (short: IR), as proposed in discursive social psychology by Potter and Wetherell ([1987] 2007). IR are recurrent linguistic units; as such, they are the manifestation of the possible legitimate (relational and complimentary) positions and interpretations within a discursive order.

To illustrate what this means, I draw on a study of the discourse of 'integration through language' (see Flubacher [2014, 2016] for more detail), which consisted of a critical analysis of its emergence in the legal framework for the (non-EU/EFTA) migrant population in Switzerland (1998–2008), which in turn resulted in the implementation of specific language requirements for granting or renewing residence and/or settlement permits. The emergence and legal development of this particular discourse is a common political attempt in Western Europe to regulate a specific contact situation resulting from migration.

For this example, I focus on an analysis of semi-standardized qualitative interviews (15 in total, conducted mostly in 2007–2008) with politicians, activists, or experts on migration and integration in the canton of Basel. Embedded in a Foucauldian perspective, the first step of the analysis (Potter and Wetherell 2007: 158–176) was to search for patterns in order to find recurring linguistic units on questions of integration and language learning in the interviews. Such units could be, for example: specific recurring formulations, specific metaphors, categories provided/created by interviewees (or co-constructed in the interviews), arguments, tropes, and/or communicative strategies. While there were several recurring patterns, two verbs soon emerged as particularly relevant, as illustrated in the following extract:

> Extract 1 (my emphasis): Head of integration, education and language (NGO)
> es gaht nid numme um (.) um (.) *fordere* sondern es gaht würklich aso vo mir här gseh gahts drum dass me denn au würklech zum könne *fördere* au alles zur verfüegig stellt was es brucht / segs jetzt sprachkurs segs integrationskürs
>
> English translation of Swiss German:
> It's not just about erm (.) erm (.) *demanding* but it's really about in my opinion it's about that in order for it to be really possible to *promote* we provide everything that is needed / be it language courses be it integration courses
>
> Transcription conventions:
> (.); (..) short pause; longer pause
> / ; \ rising intonation; falling intonation

The verbs 'demanding' (Swiss German *fordere*; German *fordern*) and 'promote' (Swiss German *fördere*; German *fördern*) are woven into this utterance as concepts that seem

to contain specific argumentative functions, which is related to the implementation of the (originally pedagogical) *leitmotif* of 'promoting and demanding' (basically, 'rights and obligations') in Swiss integration policies at the time. It is this seemingly self-evident conceptual use of the two verbs that led to their identification as the two dominant IR. In a next step, following Gilbert and Mulkay (1984), their 'basic principle' had to be extracted from the interviews (i.e. the main arguments underlying each IR): while the basic principle of the IR 'demanding' seemed to be that of enforcing language learning, the one of 'promoting' was that of supporting the migrant population in learning the local language. This can also be seen in the extract above, where the expert underscores that 'we provide everything', thereby actively taking on responsibility as a society, rather than deflecting to the individual migrants (which typically would be associated with the 'demanding' IR).

Combining the IR with positioning theory (e.g. analyzing communicative strategies of personalization or distancing; Davis and Harré 1999) clearly shows how the interviewees used the IR to position themselves, as well as (present or absent) third parties. In the extract above, the expert distances herself from 'demanding' ('it's not just about'), and instead emphasizes, and thus aligns herself with, 'promoting' ('it's really about'), which further becomes evident through the use of 'we'. Others distanced themselves in their statements from both IR, which they considered too extreme, and positioned themselves in a reasonable and moderate 'middle ground'. Finally, this short discussion illustrates already how both 'demanding' and 'promoting' are IR that are used complimentarily and relationally, all without questioning the underlying assumption that language is a requirement for integration, thus reinforcing the discursive order. The ultimate consequence of this for migrants is that linguistic assimilation has become an uncontestable requirement for them.

Summary of empirical steps:

1. Identify the most central recurring linguistic units (i.e. interpretative repertoires)
2. Identify their basic principle
3. Identify the strategic function of the IR
4. Identify the discursive order

4.2. A discursive analysis of interactions

While research interviews are interactions in their own right, infused with their own communicative and interactional rules, the following section will shift the analytical focus to another genre of interviews, i.e. the asylum interview. While maintaining a discursive lens, it will move from the level of content and argument, to the level of interaction. This means that the linguistic patterns emerging in an interaction are again understood as related to, and embedded in, particular structures of power which have consequences for that particular situation, i.e. in moments of contact between speakers with different codes, socioeconomic backgrounds, social positions, and so on. Arguing that these power structures are not necessarily enacted openly and/or consciously, Maryns

and Blommaert (2006) invoke Foucault's (1980) notion of 'capillary power', which denotes an understanding of power as pervasive in all social (and thus all linguistic) practices. Turning to interactions, they see this capillary power enacted "through habituated patterns of behaviour such as language choice, codeswitching and shifts in participation framework, none of which appear as explicit or remarkable manifestations of power" (Maryns and Blommaert 2006: 190).

This is exemplified in their discursive analysis of asylum application interviews in Dutch-speaking Belgium, in which they identify code-switching "as a contextualization cue [that] offers opportunities to regulate access to particular contextual spaces" (Maryns and Blommaert 2006: 178). In presenting several transcribed segments of a single asylum interview that involved the applicant, the interviewer, and an interpreter, they show how an applicant from Sierra Leone is conversationally excluded through practices of code-switching (Maryns and Blommaert 2006: 180–182), which they interpret as a "covert control of the discursive processes" (Maryns and Blommaert 2006: 179). To illustrate this, they segment the interview according to switches between Krio, English, and Dutch. While the interpreter is fluent in both of the languages spoken by the applicant, i.e. Krio and English, he switches to Dutch when consulting with the Belgian interviewer, which is interpreted as institutionally aligning with him. The analysis of this segment and of the rest of the interaction shows how the practice of code-switching systematically excludes the applicant from meaning-making processes directly related to his narrated account, which are crucial for establishing credibility (e.g. chronological precision and coherence) and will impact the probability of being granted asylum.

On the basis of their discursive analysis, Maryns and Blommaert (2006: 190) conclude by proposing an "awareness that language is the prime organizer of discreet, often invisible power differences in environments where the absence of power and the prevalence of rights are advertised as house culture". This erasure of power becomes especially salient in the context of European institutions that are built ideationally on the democratic pillars of equality and equity. Analyses such as these thus present an opportunity to delve beyond the ideational level in order to present a critique of "narrative inequalities" (Blommaert 2001), which have material effects on actors not equipped with knowledge of 'house culture' (see also Roberts [2013] on the 'linguistic penalty' for foreigners in job interviews in Britain). As refugees are expected to be literate, fluent in standard languages, and competent in complex institutional genres, misinterpretations at the interactional level are interpreted institutionally as narrative 'inconsistencies' that quite clearly lead to the rejection of asylum applications (Maryns 2005; see also Lehner 2018).

Summary of empirical steps:

1. Identify recurrent linguistic patterns with interactional effects
2. Identify the discursive meaning of these effects
3. Identify institutional and social consequences

4.3. A discursive analysis of language policies

This third example leads us directly to the analysis of policies, which are among the main political and institutional instruments used to regulate issues in cases of language

conflict, most importantly in the context of education, language standardization, and minority language struggles, as well as in the form of (tightening) linguistic regulations for migration and integration purposes. Again, due to the diversity of these contexts and the related policies, as well as varying research interests and disciplinary investments, a variety of analytical methods exist even within discourse analysis. In the following, I focus on Duchêne's (2008) critical sociolinguistic analysis of the discursive construction of linguistic minorities at the United Nations, which led to the *Declaration on the Rights of Persons Belonging to National or Ethnic, Religious and Linguistic Minorities* (1992). His analytical focus on historicity takes into account the production and materialization of discourses in particular discursive spaces, thereby situating them in a process of knowledge production closely connected to authority (see Foucault on archeology [1969] and genealogy [1971]). Accordingly, Duchêne was interested in why and how 'minority' becomes the discursive topic of such an institution, what the legitimate positions are, and who assumes the power of definition (see also Sokolovska [2016] for a genealogical analysis of the institutional discursive construction of 'linguistic diversity' at the Council of Europe).

For his analysis, Duchêne (2008: 39) unpacked his data as belonging to three consecutive discursive events (i.e. "key moments in history that materialize via specific documents/texts/interactions"), with each encompassing the production process of a core text (two declarations, one covenant) and the related debates. Discursive events must thus be read from a processual perspective, with one constituting the condition of possibility for the next in a veritable series. In order to understand the meaning of each discursive event in empirical practice, this ethnographically inspired approach involves the following steps (Duchêne 2008: 75–76): meticulously retracing the various drafting stages of the texts; carefully studying these drafts as well as the transcripts of the debates; and paying particular attention to the institutional mode of discourse production, i.e. the predefined textual structure, codified language, and intertextual links to preceding declarations, resolutions, and other institutional documents (cf. Urban's [1996] concept of 'entextualization').

The first analytical step is thus to identify the discursive events, which is of great help in ordering and organizing the data, as exemplified in the following. As part of the first discursive event, Duchêne (2008: 127–129) analyzes the changes from Proposition 1 to Proposition 3 regarding an article on minorities in the *Universal Declaration of Human Rights*, focusing on 'discursive movements' between the three drafts. He lists which additions, omissions, substitutions, and reformulations happened at the lexical level across the drafts (Figure 51.1). Yet as he argues, these discursive movements are more far-reaching than mere semantics. In order to understand the reasons for, and consequences of, such changes across the three proposition drafts, he emphasizes the relevance of studying critically the related debates that point to discursive constraints, and bring to the fore the underlying ideological investments and power struggles that led to such changes.

His retracing of the changes between drafts visualizes the issues that arose between the members of the Sub-Commission when attempting to define 'minority' and debating the epistemological meaning and material consequences of such a definition (e.g. the term 'minority' itself was substituted with 'groups'). The article ultimately was not included in the Declaration, precisely due to its highly contested formulation, which was by and large connected to the two then-dominant epistemes bound up with the blocs in

In States inhabited ~~by a substantial number of persons~~ of a race, language or religion other than those of the majority of the population, persons ~~belonging to such ethnic, linguistic or religious minorities~~ shall have the right as far as compatible with public order to establish and maintain their schools and cultural or religious institutions, and to use their own language in the press, in public assembly and before the courts and other authorities of the State. (E/CN.4/21 Annexe F)	In States inhabited by well defined ethnic, linguistic or religious groups which are clearly distinguished from the rest of the population and which want to be accorded differential treatment, persons belonging to such groups shall have the rights as far as is compatible with public order and security to establish and maintain their schools and cultural or religious institutions, and to use their own language and script in the press, in public assembly and before the courts and other authorities of the State, if they so choose. (E/CN.4/Sub.2/38)

Fig. 51.1: Discursive movements from Proposition 1 to Proposition 3 (Duchêne 2008: 129)

East and West. In summary, this historically embedded analysis of the institutional production process on the basis of discursive events uncovers the ideologically invested positions of the involved parties which materialize over the course of a particular institutional, discursive construction of the term 'minority'.

Summary of empirical steps:

1. Identify the relevant discursive events and the data that belong to each
2. Identify the discursive movements as well as the related arguments and consequences
3. Identify the positions
4. Identify the meaning of the positions and their discursive power

4.4. A discursive analysis of corpora

The fourth exemplary approach based on a critical Foucauldian understanding shifts the focus to discursive analysis of large corpora, i.e. 'discourse linguistic multilevel analysis' (German *diskurslinguistische Mehr-Ebenen-Analyse*, or DIMEAN; Spitzmüller and Warnke 2011a, 2011b). Spitzmüller and Warnke (2011b: 121−301) published an extensive methodological discussion of the various decisions and steps necessary to produce a coherent discourse analysis. Spitzmüller and Warnke (2011b: 135) argue against describing DIMEAN as a specific method or model, but rather view it as a methodological framework which incorporates a focus beyond the textual level for discourse analyses in linguistics. They push for a mixed method approach, most importantly when tackling a large corpus (e.g. of media), thereby combining quantitative and qualitative methods (see Spitzmüller's [2007] example of the discourse surrounding 'Anglicisms' in German media).

Central for them is that data be analyzed at several levels beyond the immediate textual surface, i.e. at the levels of the actors, the transtextual, and the intratextual. In this context, actors emerge as the articulatory hinges between the transtextual and the

intratextual levels, bringing into focus agency and practices that are couched between discursive constraints, yet also shape discursive materializations. These levels each open up the possibility of a variety of methods (for an illustrative summary, see Spitzmüller and Warnke 2011b: 201), which shall be discussed here only briefly. Analysis at the transtextual level is directed towards a discourse orientation that might focus on frames and topoi, but also on intertextuality or ideologies, indexical orders, and so on (Spitzmüller and Warnke 2011b: 187–197). The intratextual analysis is the most diversified (Spitzmüller and Warnke 2011b: 137–172), in that it can address and employ various linguistic analytical methods which concomitantly focus on either texts, propositions, or words. Finally, couched in the middle are actors and their practices, which can be analyzed (Spitzmüller and Warnke 2011b: 172–187) by taking into account their mediality (communicative patterns, forms of communication, etc.), discourse positions (with reference to concepts such as power, ideology brokers, and voice), and/or interactional roles (e.g. recipient or producer). Again, it is important to note that DIMEAN proposes the use of different methods, and does not prescribe which to draw upon; its main purpose is to address different levels, and thus to arrive at a detailed understanding of the discursive order, its effect on the production of texts, and the (potential) space for actors to maneuver.

Lehner's (2017) corpus study of online profiles of Austrian 'integration ambassadors' serves as an empirical example of how DIMEAN can be applied. The display of such profiles on a website was a concerted effort on the part of the Austrian Federal Ministry for Europe, Integration, and Foreign Affairs in 2011. In her analysis of the 113 online profiles that mention language (the corpus consisted of 276 profiles in total), Lehner pursues the question of whether these 'ambassadors', conceived as role models for other migrants, highlight their multilingual repertoire(s) as capital *sensu* Bourdieu (1977). Toward this aim, Lehner first analyzes the visual representations of the actors (i.e. the videos, statements, pictures, etc.) which were displayed in the online profiles of the 'ambassadors'. In drawing on Goffman's (1981; see Spitzmüller and Warnke 2011b: 174–176) description of the roles of actors as animators, authors, and principals, she found collusion between these roles, which pointed to a lack of transparency in the production of these profiles (Lehner 2017: 92–95), raising questions regarding voice and authority (e.g. who was selected to be an 'ambassador', who authored or revised the statements, etc.).

At the intratextual level, Lehner (2017: 96–99) analyzed frequency and quality of nouns (most importantly: *Österreich* 'Austria', *Integration* 'integration', *Wien* 'Vienna', *Sprache* 'language', and *Jahre* 'years') and verbs (interestingly, verbs of learning or knowledge acquisition, such as *lernen* 'to learn' or *erwerben* 'to acquire', occurred with high frequency), treating them as indexical of the narratives within the framework of this discourse. Finally, the transtextual analysis sought the articulation of language ideologies by comparing across texts how 'language' is framed, discussed, and narrated. Here, she assessed that a high number of 'ambassadors' wrote about an initial, difficult phase in Austria, and the (positive) consequences of language learning for them (Lehner 2017: 99–103). Her analysis of all three levels showed how multilingual competences were not considered valuable capital. This finding led her to conclude that these profiles ultimately reaffirm the discursive order of the (local) 'language as key', while simultaneously erasing other linguistic repertoires.

Summary of empirical steps:

1. Identify the relevant intratextual elements and the necessary level of detail
2. Identify the actors and their positions, roles, or mediality
3. Identify the relevant transtextual concepts, strategies, etc.
4. Identify the most suitable (linguistic) method for analysis of each level

5. Taking stock with a perspective

As described in this chapter, discourse analysis differs widely in its ontological and epistemological applications. In this contribution, the focus was on critical approaches drawing on Foucault that situate discourses as constituting (and constitutive of) social practices, as invested with ideologies and power plays, and as historically contingent. While the goal was to exemplify the methodological stakes and empirical practices of discourse analysis on the basis of four analyses which tackle different types of data, it is clear that these summaries only scratch the surface. As diverse as the interests and data are, what they have in common is an analytical focus on the materialization of discourse in debates, documents, legal texts, articles in popular media, and so on, while also seeking to investigate broader social conditions. With this in mind, the question remains what discourse analysis can tell us about social practices and what the real impacts of discourses on peoples' lives are. Although discourses and social practices are inextricably linked, how to study these links beyond the textual therefore becomes an empirical question, as has been suggested by Spitzmüller and Warnke (2011b). While studying the historicity of discourses at a textual level, e.g. in tracing the drafting process of a policy and the accompanying debates, is relevant in its own right, it does not tell us anything about how individuals and groups navigate the discourses framing their experiences, and what these discourses ultimately tell us about broader political economic conditions (see also Barakos and Unger 2016). This is all the more pressing when we recall Nelde's (1988) cautionary words for researchers: to challenge the role attributed to language in instances of contact and/or conflict; to question whether other political economic reasons, power struggles, and/or inequalities could be underlying a language conflict; to take into consideration the possibility that any so-called language conflict could actually be a symbolic conflict; and finally, to broaden the discussion as to why language so often emerges as a site of conflict, who profits from such symbolic conflicts, and what the interests lurking behind such strategies could be.

It is against the backdrop of these questions that ethnographic studies which complement discourse analysis proper have increasingly gained traction in (critical) sociolinguistics, applied linguistics, linguistic anthropology, and so on. Ethnographic researchers of contact situations and multilingual settings try to "see how language practices are connected to the very real conditions of people's lives, to discover how and why language matters to people in their own terms, and to watch processes unfold over time" (Heller 2008: 250), most commonly through participant observation and interviews, all the while attentive to the hegemonic discourses that structure and determine social action and practices. Notable examples of such attempts have emerged in the research domain addressing issues of social and structural inequality, particularly in the context of education (Martín Rojo 2010; Pérez-Milans 2013), integration and migration (Codó 2009;

Del Percio 2016; Sabate Dalmau 2016), and work or labor (Duchêne, Moyer, and Roberts 2013; Garrido 2017; Kraft 2016; Lorente 2017). Among others, these fields of research are extremely valid to the study of contact (and underlying conflicts) between speakers, as various political economic interests are at stake. In these instances, the potentially differing investments and positions of speakers materialize, while a central role is attributed to language in education, integration or 'belonging', and work. This is, in sum, why I argue for combined approaches and methods when analyzing discourses and their consequences for speakers, individuals, and social groups.

In conclusion, another point should be taken into consideration: if contact linguistics bears explanatory potential for the roles, practices, and forms of language in situations and contexts of multilingualism of various kinds, then an ethnographic discourse analysis provides a lens through which to view the broader surrounding issues at play which impact how languages and their speakers are situated in these very contexts. Last, but not least, a reflexive discourse analysis stimulates researchers to take into consideration the influence of their studies of multilingual encounters and situations, especially those which are marked by unequal structures and discriminatory mechanisms, and to ponder what kind of knowledge we produce for (and to what effect on) speakers in contact situations.

Acknowledgments

First of all, I am thankful to be included in this project, a continuation of Peter Nelde's legacy. I am also indebted to the reviewers as well as the editors for comments on an earlier version of this chapter. As is often the case, this contribution is all the better for the review. Finally, my thanks to Alexandre Duchêne, Jürgen Spitzmüller, and Sabine Lehner, who commented on and approved the final version. Many thanks also to the editorial team for the final stylistic touches.

6. References

Barakos, Elisabeth & Johann W. Unger (eds.)
2016 *Discursive approaches to language policy*. London: Palgrave Macmillan.

Blommaert, Jan
2001 Investigating narrative inequality: African asylum seekers' stories in Belgium. *Discourse & Society* 12(4). 413–449.

Blommaert, Jan
2005 *Discourse: A critical introduction*. Cambridge: Cambridge University Press.

Bourdieu, Pierre
1977 The economics of linguistic exchanges. *Social Science Information* 16. 645–668.

Busch, Brigitta
2016 Heteroglossia of survival: To have one's voice heard, to develop a voice worth hearing. *Working Papers in Urban Language & Literacies* 188. 1–12.

Cameron, Deborah
2001 *Working with spoken discourse*. London: Sage.

Codó, Eva
2009 *Immigration and bureaucratic control: Language practices in public administration.* Berlin & New York: Mouton de Gruyter.
Darquennes, Jeroen
2015 Language conflict research: A state of the art. *International Journal of the Sociology of Language* 235. 7–32.
Davis Bronwyn & Rom Harré
1999 Positioning: The discursive production of selves. *Journal for the Theory of Social Behaviour* 20(1). 44–63.
Del Percio, Alfonso
2016 The governmentality of migration: Intercultural communication and the politics of (dis)placement in Southern Europe. *Language & Communication* 51. 87–98.
Dorostkar, Niku
2014 *(Mehr-)Sprachigkeit und Lingualismus: Die diskursive Konstruktion von Sprache im Kontext nationaler und supranationaler Sprachenpolitik am Beispiel Österreichs.* Göttingen: Vienna University Press bei V&R unipress.
Duchêne, Alexandre
2008 *Ideologies across nations: The construction of linguistic minorities at the United Nations*. Berlin & New York: Mouton de Gruyter.
Duchêne, Alexandre, Melissa Moyer & Celia Roberts (eds.)
2013 *Language, migration, and social inequalities: A critical sociolinguistic perspective on institutions and work*. Bristol: Multilingual Matters.
Flubacher, Mi-Cha
2014 *Integration durch Sprache – die Sprache der Integration: Eine kritische Diskursanalyse zur Rolle der Sprache in der Schweizer und Basler Integrationspolitik 1998–2008*. Göttingen: Vienna University Press bei V&R unipress.
Flubacher, Mi-Cha
2016 On 'promoting and demanding' integration: A discursive case study of immigrant language policy in Basel. In Elisabeth Barakos & Johann W. Unger (eds.), *Discursive approaches to language policy*, 231–252. London: Palgrave Macmillan.
Foucault, Michel
1969 *L'archéologie du savoir*. Paris: Gallimard.
Foucault, Michel
1971 *L'ordre du discours*. Paris: Gallimard.
Foucault, Michel
1975 *Surveiller et punir*. Paris: Gallimard.
Foucault, Michel
1980 *Power/Knowledge: Selected interviews and other writings (1972–77)* (Colin Gordon ed.). New York: Pantheon Press.
Foucault, Michel
1981 The order of discourse. In Robert Young (ed.), *Untying the text: A post-structuralist reader*, 51–78. Boston & London: Routledge & Kegan Paul.
Garrido, Maria Rosa
2017 Multilingualism and cosmopolitanism in the construction of a humanitarian elite. *Social Semiotics* 27. 359–369.
Gilbert, G. Nigel & Michael Mulkay
1984 *Opening Pandora's box: A sociological analysis of scientists' discourse*. Cambridge: Cambridge University Press.
Goffman, Erving
1981 *Forms of talk*. Philadelphia: University of Pennsylvania Press.
Heller, Monica
2002 *Éléments d'une sociolinguistique critique*. Paris: LAL, Didier.

Heller, Monica
2008 Doing ethnography. In Li Wei & Melissa G. Moyer (eds.), *The Blackwell guide to research methods in bilingualism and multilingualism*, 249–262. Malden: Blackwell.
Kraft, Kamilla
2016 *Constructing migrant workers: Multilingualism and communication in the transnational construction site*. Oslo: University of Oslo dissertation.
Lehner, Sabine
2017 Sprachliches Kapital und "Integration": Bourdieus sprachlicher Markt revisited am Beispiel der österreichischen "Integrationsbotschafter_innen". *Wiener Linguistische Gazette* 80. 81–107.
Lehner, Sabine
2018 The discursive construction of (in)credibility: Language ideologies and intertextuality in Austrian asylum procedures. In Iman Nick (ed.), *Forensic linguistics: Asylum seekers, refugees, and immigrants*, 95–120. Wilmington: Vernon Press.
Lorente, Beatriz
2017 *Scripts of servitude: Language, labor migration, and transnational domestic work*. Bristol: Multilingual Matters.
Martín Rojo, Luisa
2010 *Constructing inequality in multilingual classrooms*. Berlin: Mouton de Gruyter.
Martín Rojo, Luisa
2017 Language and power. In Ofelia Garcia, Nelson Flores & Massimiliano Spotti (eds.), *Oxford handbook of language and society*, 78–102. Oxford: Oxford University Press.
Maryns, Katrijn
2005 Monolingual language ideologies and code choice in the Belgian asylum procedure. *Language & Communication* 25. 299–314.
Maryns, Katrijn & Jan Blommaert
2006 Conducting dissonance: Codeswitching and differential access to context in the Belgian asylum process. In Clare Mar-Molinero & Patrick Stevenson (eds.), *Language ideologies, policies, and practices: Language and the future of Europe*, 177–190. London: Palgrave Macmillan.
Nelde, Peter H.
1988 Sprachkonfliktforschung. *Folia Linguistica* 22. 73–85.
Pérez-Milans, Miguel
2013 *Urban schools and English language education in late modern China: A critical sociolinguistic ethnography*, 177–190. London: Routledge.
Potter, Edward & Margaret Wetherell
2007 [1987] *Discourse and social psychology*. Los Angeles: Sage.
Reisigl, Martin & Ruth Wodak
2009 The discourse-historical approach (DHA). In Ruth Wodak & Michael Meyer (eds.), *Methods of critical discourse analysis*, 2nd edn., 87–121. Los Angeles: Sage.
Roberts, Celia
2013 The gatekeeping of Babel: Job interviews and the linguistic penalty. In Alexandre Duchêne, Melissa Moyer & Celia Roberts (eds.), *Language, migration, and social inequalities: A critical sociolinguistic perspective on institutions and work*, 81–94. Bristol: Multilingual Matters.
Sabate Dalmau, Maria
2016 Migrants' alternative multi-lingua franca spaces as emergent re-producers of exclusionary monolingual nation-state regimes. *Multilingua* 35(6). 649–673.
Schiffrin, Deborah
1994 *Approaches to discourse*. Oxford: Blackwell.

Schiffrin, Deborah
1996 Discourse analysis and pragmatics. In Hans Goebl, Peter H. Nelde, Zdenek Stary & Wolfgang Wölck (eds.), *Kontaktlinguistik / Contact linguistics / Linguistique de contact: Ein internationales Handbuch zeitgenössischer Forschung / An international handbook on contemporary research / Manuel international des recherches contemporaines*, 138–144. Berlin: De Gruyter.
Sokolovska, Zorana
2016 Imagining Europe's linguistic diversity in the Council of Europe's Parliamentary Assembly. *Language & Communication* 51. 40–49.
Spitzmüller, Jürgen
2007 Staking the claims of identity: Purism, linguistics and the media in post-1990 Germany. *Journal of Sociolinguistics* 11(2). 261–285.
Spitzmüller, Jürgen & Ingo H. Warnke
2011a Discourse as a linguistic object: Methodical and methodological delimitations. *Critical Discourse Studies* 8(2). 75–94.
Spitzmüller, Jürgen & Ingo H. Warnke
2011b *Diskurslinguistik: Eine Einführung in Theorien und Methoden der transtextuellen Sprachanalyse*. Berlin: De Gruyter.
Urban, Greg
1996 Entextualization, replication, and power. In Michael Silverstein & Greg Urban (eds.), *Natural histories of discourse*, 21–44. Chicago: University of Chicago Press.
Wodak, Ruth & Michael Meyer
2009a Critical discourse analysis: History, agenda, theory and methodology. In Ruth Wodak & Michael Meyer (eds.), *Methods of critical discourse analysis*, 2nd edn., 1–33. Los Angeles: Sage.
Wodak, Ruth & Michael Meyer
2009b *Methods of critical discourse analysis*, 2nd edn. Los Angeles: Sage.

Mi-Cha Flubacher, Vienna (Austria)

52. Corpus linguistic methods

1. Introduction
2. Overview of existing corpora and their relevance to contact linguistics
3. Transcription and annotation tools
4. Corpus-driven studies in contact linguistics
5. Conclusion
6. References

1. Introduction

Corpus linguistics is a methodological approach that takes an empirical stance toward the study of language. It relies on the analysis – whether qualitative or quantitative – of a body of written texts and/or transcriptions of spontaneous and/or semi-spontaneous speech. Corpus linguistic methods potentially have a strong impact on theory, as they can offer support for (or challenge) theoretical assumptions. They can also help improve

https://doi.org/10.1515/9783110435351-052

our understanding of previously described linguistic phenomena, and can reveal new areas of investigation that had hitherto gone unnoticed. Moreover, corpus linguistics is closely related to various fields of applied linguistics, for example, through the elaboration of pedagogical tools.

More specifically, in contact linguistics, a corpus-driven approach based on ecologically valid data allows for the examination of the constraints and social significance of bilingual speech. In addition, in experimental approaches to bilingualism, natural corpus data are used as basic frequency data in conjunction with controlled data produced in a laboratory environment (see Chapter 53 for details).

2. Overview of existing corpora and their relevance to contact linguistics

The use of corpus linguistic methods in the various subfields of linguistics depends largely on practical issues related to the availability of corpora and/or the ease with which a corpus can be built and searched. This section presents a brief overview of existing corpora, and the ways in which they can be explored for language contact. Readers should be aware that this and subsequent sections of the chapter contain no hyperlinks to corpora or corpus linguistic tools in the running text; such links can be found in the references (Section 6).

2.1. Written corpora

Written corpora have become fairly easy to compile, thanks to widespread use of computers and the internet. As a result, large electronic corpora of formal written speech, often comprising literary and journalistic texts, are currently available online in a variety of languages. For example, several million-word corpora exist even for small national languages (such as Albanian) and minority languages (such as Basque and Soviet Romani). Following the model of well-developed English corpora, electronic corpora aim to provide users with tools to conduct online and/or offline searches, not only for word frequency, but also for collocations, concordances, and *n*-grams.

Despite this considerable progress in corpus compilation, sizeable corpora from a variety of speakers and sources are still restricted to a small number of major languages of communication. For example, Anand, Chung, and Wagers (2011) observe that 85 % of the data that are part of the Linguistic Data Consortium (www.ldc.upenn.edu) represent only five languages, i.e. English, Chinese, Arabic, Spanish, and Japanese.

From a contact linguistic perspective, language contact phenomena are relatively limited as found in the formal written texts of national languages. In contrast, informal written texts, compiled for example from social media posts and text messages (SMS), may provide information about a wider range of contact phenomena, such as code-switching.

2.2. Spoken corpora

Spoken corpora are of special interest to linguists focusing on contact, as they potentially include a wide range of contact phenomena and may reveal information about the cognitive processes involved in production and comprehension of bilingual speech. Large spoken corpora of adult language have been constituted on the basis of scripted and/or unscripted speech collected from television and radio broadcasting (a case in point is the British National Corpus). Some smaller corpora provide transcripts and recordings of free-speech conversations (e.g. the Santa Barbara Corpus of Spoken American English and the Switchboard Corpus). Most of these corpora, however, do not target interactions between bilinguals, and can at best provide evidence for the study of borrowings.

Linguists interested in bilingualism and its short- and long-term effects have compiled new corpora involving adult bilingual speakers, whether from (post)colonial settings, traditionally bilingual communities, or migrant communities. In order to document the most ordinary ways of speaking, recordings of unscripted speech were made either in natural conditions, or during interviews (preferably conducted by community members). These recordings are combined with corpora from scripted speech, and from transcripts of television and radio broadcasting. However, we note that, in most cases, these corpora are not made available to members of the scientific community.

2.3. Spoken corpora from endangered languages

A promising contribution of corpus linguistics to the study of language contact phenomena may be its growing use in language documentation initiatives. Indeed, corpora from endangered languages are increasingly compiled and made available online. Examples include the Endangered Languages Archive (ELAR), the Documentation of Endangered Languages Archive (DoBeS), the Pacific and Regional Archive for Digital Sources in Endangered Cultures Archive (PaRADiSEC), the Archive of the Indigenous Languages of Latin America (AILLA) collection, and the Pangloss Collection.

2.4. Sign language corpora

In past decades, research on sign languages has increased, typically relying on video corpora of signed interactions (e.g. the Sign Language corpus at the Language Archive). As summarized in Zeshan et al. (in press), there are three types of research studies on bilingual sign language: (i) research on sign bilingualism of bilinguals who use a sign language and a spoken language in its written form; (ii) research on bimodal bilingualism of hearing individuals who use a sign language and a spoken language in its spoken form; and (iii) research on code-switching of bilinguals who use more than one sign language.

2.5. Bilingual acquisition corpora

Spoken language corpora have held a privileged position in both first and second language acquisition studies. Bilingual acquisition has been explored through methods ran-

ging from parental diaries to extensive recordings of natural speech. Examples can be found in the Child Language Data Exchange System (CHILDES).

2.6. Learner corpora

Spoken or written learner corpora, such as the International Corpus of Learner English (ICLE), are increasingly used to explore the effects of an L1 on an L2 (or L3). Learner corpus research not only contributes to theories of second language acquisition, but also plays an important role in applied linguistics through the elaboration of new pedagogical tools (see Granger, Gilquin, and Meunier 2015).

2.7. Parallel corpora

The term 'multilingual corpora' covers what is also known as 'parallel corpora', which consist of a body of texts translated into different languages aligned across languages, e.g. the European Parliament Proceedings Parallel Corpus 1996–2011. Parallel corpora are mainly used for research in contrastive linguistics and translation studies. They also have a great impact in translation training and the development of translation tools for language learners. The Multilingual Student Translation Corpus (started in 2016) aims to compile a large, new multilingual student translation corpus that will bring together these two research paradigms, i.e. learner corpus research and corpus-based translation studies.

3. Transcription and annotation tools

In contact linguistics, corpus-based research is a rapidly growing field, yet the development of bilingual corpora is associated with unique challenges. Indeed, despite considerable progress in the development of automated techniques, bilingual corpora remain extremely challenging for natural language processing technologies, i.e. automatic speech recognition, parsing, machine translation, information retrieval and extraction, and computational processing (see Çetinoğlu, Schultz, and Vu 2016; Guzmán et al. 2016; Partanen et al. 2018; Zhang et al. 2018). This means that manual coding is still needed in order to resolve issues that arise from the competing grammars and phonologies of the languages in use. As a result, building a large bilingual corpus is very time consuming.

It therefore becomes urgent for linguists to embrace collaborative approaches to linguistic data compilation, for example, through online annotation of open data, including through crowdsourcing (see Section 6.2.2. for a list of examples). Such initiatives can be encouraged by taking into consideration the potential role of creating a corpus (based on a set of quality standards and citations) in the career of a linguist (see discussion in Thieberger et al. 2015). Of course, providing access to raw and/or annotated data on which corpus-based research relies raises ethical questions. In order to make primary and/or annotated data accessible to the scientific community while still ensuring the pro-

tection of personal data, one must ensure that the privacy of the speakers is respected (for privacy regulations in Europe, see ec.europa.eu/digital-single-market/en/open-access-scientific-knowledge-0; for privacy regulations in direct relation to language resources, see www.clarin-d.de/en/help/legal-information-platform). That can be done through anonymizing the portions of sound files and/or transcripts that may include personal information.

From the perspective of open-access data, the annotation standards (Section 6.2.1) used for a corpus gain in significance. The Worldwide Web Consortium (W3C; www.w3.org/Consortium/) recommends the Extensible Markup Language (XML) and Text Encoding Initiative (TEI) standards for annotation standards. For example, in an XML file, content is delimited by tags and surrounded by angle brackets in order to describe a document and specify information about it, i.e. <sentence>, in angle brackets, marks the beginning of a sentence and </sentence>, in angle brackets with a slash, marks its end. Depending on the research questions, annotators may provide an orthographic transcription, parse the corpus for syntactic analysis, and annotate parts of speech (nouns, proper nouns, verbs, conjunctions, particles, etc.) by using, for example, Universal Parts Of Speech (POS) tags. For lesser-known languages, a morpheme-by-morpheme transcription and translation (or *interlinearization*) is often provided together with the free translation at the sentence level. These levels (i.e. words, morphemes, glosses) should also be delimited in XML using the aforementioned tags. Within these tags, one can add information about the language, e.g. xml:lang="fr" for French. For bilingual corpora, in particular, one could tag the words depending on the languages involved. One could, e.g. use tags such as L1 (the word is used in monolingual or bilingual contexts, but can be clearly recognized in this context as part of the L1), L2 (the word is used in monolingual or bilingual contexts, but can be clearly recognized in this context as part of the L2), multiple or ambiguous (words in this context could be part of either the L1 or the L2), mixed (the word is partially in both languages, but does not exist as such in either the L1 or the L2), foreign words (for words from an L3 or past-contact languages), unknown (the annotator(s) cannot determine the language and meaning of the word), or other (non-words, or in written corpora, symbols, emoticons, etc.). For a speech corpus, one can also provide a phonetic and phonological transcription, as well as a prosodic annotation, for example, by segmenting the corpus into intonation units. The use of UNICODE (e.g. UTF-8) is strongly recommended, as it guarantees the recognition of all characters on the Internet and on any computer. In XML, one can also tag the beginning and the end of an associated sound and video file (*time alignment*) by using angle brackets and adding the relevant content information, e.g. <AUDIO start="0.1106" end="3.7594"/>. Concerning sound files, WAV format is preferred for long-term archiving, whereas MP3 format is a good option for online access.

Moreover, a vital step in the compilation of a corpus is the production and maintenance of the associated metadata. Metadata may include sociolinguistic information about the speakers, the researchers, the content of the document, and the date and place of recording, allowing for extralinguistic factors to be analyzed. Metadata may follow the Dublin Core specifications as defined by the Open Language Archives Community (OLAC), or use fuller standards, such as the ISLE Metadata Initiative (IMDI) developed by the Max Plank Institute, and the Component MetaData Infrastructure (CMDI) initiated by CLARIN and linked to the CLARIN Concept Registry. Arbil is a practical metadata editor for both IMDI and CMDI standards (Withers 2012).

Researchers should opt for annotation tools that conform to these standards and are compatible with one another. A variety of general, manual annotation tools can be adapted to the study of language contact phenomena. A very popular tool is ELAN, which synchronizes the audio and/or video files with annotation tiers (Sloetjes and Wittenburg 2008). Language contact phenomena can be tagged in these tiers, depending on the researchers' needs, and searches can be conducted with ELAN's concordance tool. Besides ELAN, ANVIL is often used in sign language annotation (Kipp 2014). EXMARaLDA is another widespread annotation tool that has been tested for multilingual research (Schmidt and Wörner 2014). CLAN and CHAT are toolsets native to language files in the CHILDES database. PHON is its phonological software program, which allows comparisons between targeted and produced phonological forms.

For phonetics, EasyAlign is used to semiautomatically align phonetic annotations with transcriptions in Praat. The WebMAUS service of the Bavarian Archive for Speech Signals (BAS; Kisler, Schiel, and Sloetjes 2012) is another tool that has successfully been tested for forced alignment. A similar tool is SailAlign, which can accommodate long and/or noisy data as well as transcription errors.

More information on corpus linguistic methods and best practices is available in Biber and Reppen (2015), Durand, Gut, and Kristofferson (2014), Gries and Berez (2017), Kirk and Andersen (2016), Lüdeling and Kytö (2008, 2009), and O'Keeffe and McCarthy (2010).

4. Corpus-driven studies in contact linguistics

4.1. Brief overview

Corpora are currently exploited in contact linguistics both in a top-down and bottom-up fashion. Top-down approaches rely on theoretical models that were primarily elaborated with monolingual data, and aim to test the validity of these models for bilingual data. This is the case for the Matrix Language Frame (MLF) model (Myers-Scotton and Jake 2017), which builds on the Speaking model in psycholinguistics (Levelt 1989) and the Minimalist approach to code-switching (MacSwan 2016), which considers that the Minimalist Program (Chomsky 1995) should also successfully account for bilingual data. Corpus studies can test the generalizations and predictions made by these models. For example, Herring et al. (2010) and Parafita-Couto and Gullberg (2017) investigate code-switches occurring between determiners and their nouns in Spanish-English, Welsh-English, and Papiamento-Dutch bilingual speech in order to evaluate competing theories of code-switching, namely the MLF model and generative approaches. On the other hand, bottom-up approaches consider that higher levels of abstraction are shaped by language practices. Linguists in this line of research typically observe patterns in bilingual corpora and aim to explain these patterns by paying close attention to factors such as frequency and priming, as with variationist approaches (Poplack 1980; Poplack and Dion 2012; Torres Cacoullos and Travis 2016) and usage-based approaches (Backus 1992, 2015; Quick et al. 2017).

Corpora can be used in both a corpus-illustrated approach (i.e. to provide occurrences of specific phenomena) and a quantitative approach (i.e. to provide frequency informa-

Tab. 52.1: Brief overview of corpus-driven studies in contact linguistics

Type of contact	Language pairs	Locations	References
Postcolonial and colonial settings	English-Spanish	USA (New York, Los Angeles, New Mexico)	Poplack 1980; Silva-Corvalán 1986; Torres Cacoullos and Travis 2016
	Quechua-Spanish	Bolivia, Ecuador	Bakker and Hekking 2012; Gomez Rendon 2008; van Hout and Muysken 1994
	Guaraní-Spanish	Paraguay	Gomez Rendon 2008
	Otomí-Spanish	Mexico	Gomez Rendon 2008
	Swahili-English	Kenya	Myers-Scotton 1993
	Hindi-English	India	Si 2010
	Chinese-English	mainland China and Hong Kong	Wang and Liu 2013
	Welsh-English	Wales	Deuchar 2006
	French-English	Canada	Poplack and Dion 2012
	Arabic-French	Morocco	Nait M'Barek and Sankoff 1988
	Nkep (Oceanic)-Bislama (English lexifier creole)	Vanuatu	Meyerhoff 2014
	Aboriginal English/ Kriol-Gurindji (Pama-Nyungan)	Australia	McConvell and Meakins 2005
Immigrants and immigrant communities	Turkish-Dutch	Netherlands	Backus 1992
	Romani-Finnish and Romani-Turkish-Greek	Greece, Finland	Adamou and Granqvist 2015
Traditional languages	Bora (Huitotoan)-Resígaro (Arawakan)	Colombian-Peruvian Amazon region	Seifart 2015

tion as well as comparisons between speakers and texts). For quantitative analyses, appropriate statistical models need to be used, depending on, among other things, the size of the corpora (see Gries 2015; Newman, Baayen, and Rice 2011; Nock et al. 2009; Tagliamonte and Baayen 2012).

At present, corpus-driven studies in contact linguistics rely on data from three main types of contact settings: postcolonial and colonial settings, immigrant communities, and traditional contact settings involving two or more native languages. Table 52.1 presents a brief overview of such studies, and shows that most of them focus on language pairs that include a European language, such as Spanish, English, and French. Studies on language contact among immigrants and immigrant communities constitute another dy-

namic area of research in corpus linguistics. In contrast, contact involving traditional languages seems to be the least explored area in corpus linguistic studies.

4.2. Case studies

Once a corpus has been collected and annotated, researchers may conduct a number of searches depending on their research questions and the specificities of the corpus. Adamou (2016) and Guzmán et al. (2016) suggest that comparable code-switching data are needed in order to understand how frequent code-switching really is in natural human communication. Based on 17 bi- or multilingual corpora, Adamou (2016) observes that most corpora contain less than 5 % word tokens from the L2, and more rarely, up to 20–35 % L2 word tokens (see Figure 52.1). These frequencies may also shed light on the language switching costs that have been noted in the experimental literature. Recently, Adamou and Shen (2019) and Johns, Valdés Kroff, and Dussias (2018) showed that language switching costs align with the code-switching habits of the community as documented in corpus studies.

The typological scale of language mixing presented in Figure 52.1 needs to be enriched by other layers of annotation. One topic of investigation is the degree of convergence between two languages in contact, and its relationship to code-switching. Torres Cacoullos and Travis (2016), for example, present new corpus data from Spanish-English speakers from a traditionally bilingual community in New Mexico in the US. They show that, despite frequent code-switching, pro-drop in the bilingual data is comparable to pro-drop in the monolingual varieties of Spanish and English. The reverse seems to be



Fig. 52.1: A scale of language mixing based on bilingual corpora (adapted from Adamou 2016)

true for the Colloquial Upper Sorbian (Slavic)-German bilinguals, who do not engage in frequent code-switching with German, and yet their pro-drop Slavic language has converged with non-pro-drop German (Adamou et al. 2016).

Pausing behavior in relation to code-switching is another research question that can only be addressed by making use of spoken corpora. Gardner-Chloros, McEntee-Atalianis, and Paraskeva (2013) recently showed that pauses do not correlate with code-switching. They came to this conclusion on the basis of a study of two bilingual Greek Cypriot-English communities with different code-switching habits: Cypriots in Great Britain, for whom code-switching is the default mode of communication, and Greek Cypriot speakers in Cyprus, for whom code-switching is a conversationally marked mode.

The existence of community-level patterns in code-switching is another interesting topic that can only be tested with naturalistic data. Gardner-Chloros (2009), for example, notes that code-switching is characterized by high variability across individuals. Adamou (2016), however, suggests that small bilingual communities may develop conventionalized ways of code-switching. Code-switching patterns were also noted in a study based on Bollywood scripts (Si 2010). This study shows that alternational code-switching between Hindi and English has increased considerably in past decades, as compared to previous code-switching patterns that relied mainly on single-word insertions.

Finally, corpus studies allow us to move beyond anecdotal evidence for so-called 'typologically-rare phenomena'. Seifart (2015), for example, draws attention to previously unreported cases of affix borrowing in two Amazonian languages, without any borrowing of the loanwords containing those affixes. Similarly, Romani corpora from Greece and Finland illustrate the regular insertion of L2 verbs with L2 verb morphology into L1-dominant speech (Adamou and Granqvist 2015), whereas in most code-switching corpora, contact verbs are regularly integrated into the morphology of the dominant language.

5. Conclusion

Corpus-driven research is a very dynamic paradigm in contact linguistics, as in other subfields of linguistics. This dynamism is due to the technological progress that have enabled the growth of computerized corpora as well as the possibilities of online access to corpora, tools, and annotation platforms. Corpus linguistics is also particularly successful because it adequately addresses public and political concerns about increased accountability and replicability in science, and also fits the requirements of open science policies. In addition, corpus linguistics provides a solid empirical basis on which to test and shape theoretical models.

However, progress must still be made. The compilation of bi- and multilingual corpora remains a particularly time-consuming task, as it requires considerable amounts of manual coding. Collaboration with Natural Language Processing specialists is therefore crucial for the development of new automated techniques adapted to the specificities of bi- and multilingual corpora. In the long-term, increasing institutional support for FAIR data (i.e. Findable, Accessible, Interoperable, and Reusable data) and the development of web-based crowdsourcing will undoubtedly help boost the production of bi- and multilingual corpora. The analysis of larger bi- and multilingual corpora from a wider variety of contact settings could in turn improve our understanding of this important aspect of human communication.

6. References

6.1. Literature

Adamou, Evangelia
2016 *A corpus-driven approach to language contact: Endangered languages in a comparative perspective*. Boston & Berlin: Mouton de Gruyter.

Adamou, Evangelia & Kimmo Granqvist
2015 Unevenly mixed Romani languages. *International Journal of Bilingualism* 19(5). 525–547.

Adamou, Evangelia & Xingjia Rachel Shen
2019 There are no language switching costs when codeswitching is frequent. *International Journal of Bilingualism* 23(1). 53–70.

Adamou, Evangelia, Walter Breu, Lenka Scholze & Xingjia Rachel Shen
2016 Borrowing and contact intensity: A corpus-driven approach from four Slavic minority languages. *Journal of Language Contact* 9(13). 515–544.

Anand, Pranav, Sandra Chung & Matthew Wagers
2011 Widening the Net: Challenges for gathering linguistic data in the digital age. Submitted to the National Science Foundation SBE 2020 planning activity. https://www.nsf.gov/sbe/sbe_2020/2020_pdfs/Wagers_Matthew_121.pdf. (accessed on 16 January 2017).

Backus, Ad
1992 *Patterns of language-mixing: A study in Turkish-Dutch bilingualism*. Wiesbaden: Harrassowitz.

Backus, Ad
2015 A usage-based approach to codeswitching: The need for reconciling structure and function. In Gerald Stell & Kofi Yakpo (eds.), *Code-switching between structural and sociolinguistic perspectives*, 19–37. Berlin: Mouton de Gruyter.

Bakker, Dik & Ewald Hekking
2012 Constraints on morphological borrowing: Evidence from Latin America. In Lars Johanson & Martine Robbeets (eds.), *Copies vs. cognates in bound morphology*, 187–220. Leiden: Brill.

Biber, Douglas & Randi Reppen (eds.)
2015 *The Cambridge handbook of English corpus linguistics*. Cambridge: Cambridge University Press.

Çetinoğlu, Özlem, Sarah Schultz & Thang Vu
2016 Challenges of computational processing of code-switching. In Mona Diab, Pascale Fung, Mahmoud Ghoneim, Julia Hirschberg & Thamar Solorio (eds.), *Proceedings of the Second Workshop on Computational Approaches to Code Switching*, 1–11. Association for Computational Linguistics.

Chomsky, Noam
1995 *The minimalist program*. Cambridge, MA: MIT Press.

Deuchar, Margaret
2006 Welsh-English code-switching and the matrix language frame model. *Lingua* 116(11). 1986–2011.

Durand, Jacques, Ulrike Gut & Gjert Kristofferson (eds.)
2014 *The Oxford handbook of corpus phonology*. Oxford: Oxford University Press.

Gardner-Chloros, Penelope
2009 *Code-switching*. Cambridge: Cambridge University Press.

Gardner-Chloros, Penelope, Lisa McEntee-Atalianis & Marilena Paraskeva
2013 Codeswitching and pausing: An interdisciplinary study. *International Journal of Multilingualism* 10(1). 1–26.

Gomez Rendon, Jorge A.
2008 *Typological and social constraints on language contact: Amerindian languages in contact with Spanish*. Utrecht: LOT.
Granger, Sylviane, Gaëtanelle Gilquin & Fanny Meunier (eds.)
2015 *The Cambridge handbook of learner corpus research*. Cambridge: Cambridge University Press.
Gries, Stefan Th.
2015 Some current quantitative problems in corpus linguistics and a sketch of some solutions. *Language and Linguistics* 16(1). 93–117.
Gries, Stefan Th. & Andrea Berez
2017 Linguistic annotation in/for corpus linguistics. In Nancy Ide & James Pustejovsky (eds.), *Handbook of linguistic annotation*, 379–409. Berlin & New York: Springer.
Guzmán, Gualberto A., Jacqueline Serigos, Barbara E. Bullock & Almeida J. Toribio
2016 Simple tools for exploring variation in code-switching for linguists. In Mona Diab, Pascale Fung, Mahmoud Ghoneim, Julia Hirschberg & Thamar Solorio (eds.), *Proceedings of the second workshop on computational approaches to code switching*, 12–20. Association for Computational Linguistics.
Herring, John R., Margaret Deuchar, M. Carmen Parafita Couto & Monica Moro Quintanilla
2010 'I saw the madre': Evaluating predictions about codeswitched determiner-noun sequences using Spanish-English and Welsh-English data. *International Journal of Bilingual Education and Bilingualism* 13. 553–573.
Johns, Michael A., Jorge R. Valdés Kroff & Paola E. Dussias
2018 Mixing things up: How blocking and mixing affect the processing of codemixed sentences. *International Journal of Bilingualism.* doi/10.1177/1367006917752570.
Kipp, Michael
2014 ANVIL: A Universal Video Research Tool. In Jacques Durand, Ulrike Gut & Gjert Kristofferson (eds.), *The Oxford handbook of corpus phonology*, 420–436. Oxford: Oxford University Press.
Kirk, John M. & Gisle Andersen (eds.)
2016 Compilation, transcription, markup and annotation of spoken corpora. [Special issue]. *International Journal of Corpus Linguistics* 21(3).
Kisler, Thomas, Florian Schiel & Han Sloetjes
2012 Signal processing via web services: The use case WebMAUS. *Proceedings of Digital Humanities 2012*, 30–34.
Levelt, Willem J. M.
1989 *Speaking, from intention to articulation*. Cambridge, MA: MIT Press.
Lüdeling, Anke & Merja Kytö (eds.)
2008 *Corpus linguistics: An international handbook*, vol. 1. Berlin: Mouton de Gruyter.
Lüdeling, Anke & Merja Kytö (eds.)
2009 *Corpus linguistics: An international handbook*, vol. 2. Berlin: Mouton de Gruyter.
MacSwan, Jeff
2016 Codeswitching in adulthood. In Elena Nicoladis & Simona Montanari (eds.), *Lifespan perspectives on bilingualism*, 183–200. Berlin: Mouton de Gruyter and Washington, DC: American Psychological Association.
McConvell, Patrick & Felicity Meakins
2005 Gurindji Kriol: A mixed language emerges from code-switching. *Australian Journal of Linguistics* 25(1). 9–30.
Meyerhoff, Miriam
2014 Borrowing in apparent time: With some comments on attitudes and universals. *Selected Papers from NWAV 42: University of Pennsylvania Working Papers in Linguistics* 20(2). 121–128.

Myers-Scotton, Carol
1993 *Duelling languages: Grammatical structure in code-switching*. Oxford: Clarendon press.
Myers-Scotton, Carol & Janice Jake
2017 Revisiting the 4-M model: Codeswitching and morpheme election at the abstract level. *International Journal of Bilingualism* 21(3). 340–366.
Naït M'Barek, Mohammed & David Sankoff
1988 Le discours mixte arabe-français: Des emprunts ou des alternances de langue? *Revue Canadienne de Linguistique* 33. 143–154.
Newman, John, Harald R. Baayen & Sally Rice (eds.)
2011 *Corpus-based studies in language use, language learning, and language documentation.* Amsterdam: Rodopi.
Nock, Richard, Pascal Vaillant, Claudia Henry & Frank Nielsen
2009 Soft memberships for spectral clustering, with application to permeable language distinction. *Pattern Recognition* 42. 43–53.
O'Keeffe, Anne & Michael McCarthy (eds.)
2010 *The Routledge handbook of corpus linguistics*. London: Routledge.
Parafita-Couto, M. Carmen & Marianne Gullberg
2017 Code-switching within the noun phrase: Evidence from three corpora. *International Journal of Bilingualism*. https://doi.org/10.1177/1367006917729543.
Partanen, Niko, Lim Kyung Tae, Michael Riessler & Thierry Poibeau
2018 Dependency parsing of code-switching data with cross-lingual feature representations. In Tommi A. Pirinen, Michael Rießler, Jack Rueter, Trond Trosterud, Francis M. Tyers (eds.), *Proceedings of the fourth international workshop on computational linguistics of Uralic languages*, 1–17. Helsinki.
Poplack, Shana
1980 Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics* 18(7). 581–618.
Poplack, Shana & Nathalie Dion
2012 Myths and facts about loanword development. *Language Variation and Change* 24(3). 279–315.
Quick, Antje Endesfelder, Elena Lieven, Malinda Carpenter & Michael Tomasello
2017 Identifying partially schematic units in the code-mixing of an English and German speaking child. *Linguistic Approaches to Bilingualism* 8(4). 477–501.
Schmidt, Thomas & Kai Wörner
2014 EXMARaLDA. In Jacques Durand, Ulrike Gut & Gjert Kristofferson (eds.), *Handbook of corpus phonology*, 402–419. Oxford: Oxford University Press.
Seifart, Frank
2015 Direct and indirect affix borrowing. *Language* 91(3). 511–531.
Sloetjes, Han & Peter Wittenburg
2008 Annotation by category: ELAN and ISO DCR. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis & Daniel Tapias (eds.), *Proceedings of the sixth international conference on language resources and evaluation*, 816–820. Marrakech.
Si, Aung
2010 A diachronic investigation of Hindi-English code-switching, using Bollywood film scripts. *International Journal of Bilingualism* 15(4). 388–407.
Silva-Corvalán, Carmen
1986 Bilingualism and language change: The extension of *estar* in Los Angeles Spanish. *Language* 62(3). 587–608.
Tagliamonte, Sali & Harald R. Baayen
2012 Models, forests, and trees of York English: was-were variation as a case study for statistical practice. *Language Variation and Change* 24(2). 135–178.

Thieberger, Nick, Anna Margetts, Stephen Morey & Simon Musgrave
2015 Assessing annotated corpora as research output. *Australian Journal of Linguistics* 36(1). 1–21.
Torres Cacoullos, Rena & Catherine Travis
2016 Two languages, one effect: Structural priming in spontaneous code-switching. *Bilingualism: Language and Cognition* 19(4). 733–753.
van Hout, Roeland & Pieter Muysken
1994 Modelling lexical borrowability. *Language Variation and Change* 6(1). 39–62.
Wang, Lin & Haitao Liu
2013 Syntactic variations in Chinese-English code-switching. *Lingua* 123(1). 58–73.
Withers, Peter
2012 Metadata management with Arbil. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the eighth international conference on language resources and evaluation*, 72–75. Istanbul.
Zhang, Yuan, Jason Riesa, Daniel Gillick, Anton Bakalov, Jason Baldridge & David Weiss
2018 A fast, compact, accurate model for language identification of codemixed text. In Ellen Riloff, David Chiang, Julia Hockenmaier & Junichi Tsujii (eds.), *Proceedings of the 2018 conference on empirical methods in natural language processing*, 328–337. Brussels.
Zeshan, Ulrike, Jenny Webster, Onno Crasborn & Evangelia Adamou
In press Forces shaping sign multilingualism. In Ulrike Zeshan & Jenny Webster (eds.), *Sign multilingualism*. Berlin: De Gruyter Mouton & Nijmegen: Ishara Press.

6.2. Online collaborative platforms, corpora and tools (selection)

last access (for all the websites listed below): 4 April 2019

6.2.1. Annotation standards

Arbil
: A general metadata editor, browser & organizer tool for IMDI, CMDI and similar metadata formats. http://tla.mpi.nl/tools/tla-tools/arbil/

Clarin
: A research infrastructure that was initiated from the vision that all digital language resources and tools from all over Europe and beyond are accessible through a single sign-on online environment for the support of researchers in the humanities and social sciences. http://www.clarin.eu/content/component-metadata

Data Category Repository (DCR), known as ISOCat
: https://openskos.meertens.knaw.nl/ccr/browser/

Dublin Core Metadata Initiative
: An open organization supporting innovation in metadata design and best practices across the metadata ecology. http://dublincore.org

TalkBank
: Constructs sample databases within each of the subfields studying communication and uses them to advance the development of standards and tools for creating, sharing, searching, and commenting upon primary materials via networked computers. https://talkbank.org/

Text Encoding Initiative (TEI)
: A consortium that develops and maintains annotation standards. http://www.tei-c.org/index.xml

Unicode
: A consortium that enables people around the world to use computers in any language. http://www.unicode.org/

Universal Dependencies
: A framework for cross-linguistically consistent grammatical annotation. http://universaldependencies.org

6.2.2. Collaborative platforms and tools

Annis
: A web browser-based search and visualization architecture for complex multilayer linguistic corpora with diverse types of annotation. http://corpus-tools.org/annis

ANVIL
: A video annotation research tool. http://www.anvil-software.org/

Atomic
: A cross-platform multi-layer corpus annotation tool. http://corpus-tools.org/atomic/

BAS
: Bavarian Archive for Speech Signals. http://clarin.phonetik.uni-muenchen.de/BASWebServices/

Brat
: Online environment for collaborative text annotation. http://brat.nlplab.org/

EasyAlign
: A user-friendly automatic phonetic alignment tool for continuous speech under Praat. http://latlcui.unige.ch/phonetique/easyalign.php

ELAN
: A professional tool for the creation of complex annotations on video and audio resources, created by the Max Planck Institute for Psycholinguistics (Nijmegen, The Netherlands). https://tla.mpi.nl/tools/tla-tools/elan/

EXMARaLDA
: A system for working with oral corpora on a computer. http://exmaralda.org/en/

IMS Corpus Workbench (CWB)
: A collection of open-source tools for managing and querying large text corpora (ranging from 10 million to 2 billion words) with linguistic annotations. http://cwb.sourceforge.net/

Praat
: Doing phonetics by computer. http://www.fon.hum.uva.nl/praat/

SailAlign
: A software toolkit for robust long speech-text alignment. http://sail.usc.edu/old/software/SailAlign/

Webanno
: A flexible, web-based and visually supported system for distributed annotations. https://webanno.github.io/webanno/

6.2.3. Corpora

Afro-Asiatic languages

Corpoafroas
: http://corpafroas.huma-num.fr/Archives/ListeFichiersELAN.php

Albanian

Albanian National Corpus
http://web-corpora.net/AlbanianCorpus/search/

Basque

Corpus of Contemporary Basque
https://www.ehu.eus/en/web/eins/egungo-testuen-corpusa-etc-

Child Language

Child Language Data Exchange System (CHILDES)
https://childes.talkbank.org

Endangered languages

Endangered Languages Archive (ELAR)
https://elar.soas.ac.uk/
Documentation of Endangered Languages (DoBeS)
http://dobes.mpi.nl/
Pacific and Regional Archive for Digital Sources in Endangered Cultures (PaRADiSEC)
http://www.paradisec.org.au/home.html
Pangloss Collection. An archive for endangered languages
http://lacito.vjf.cnrs.fr/pangloss/index_en.html

English

British National Corpus
http://www.natcorp.ox.ac.uk/
Corpus Brigham Young
https://corpus.byu.edu/ (groups the most widely used online corpora of English)
Corpora of the Centre for English Corpus Linguistics (Louvain-la-Neuve, BE)
https://uclouvain.be/en/research-institutes/ilc/cecl/corpora.html
International Corpus of Learner English (ICLE)
https://uclouvain.be/en/research-institutes/ilc/cecl/icle.html
Santa Barbara Corpus of Spoken American English
http://www.linguistics.ucsb.edu/research/santa-barbara-corpus

European Union's official languages

European Parliament Proceedings Parallel Corpus 1996–2011
http://www.statmt.org/europarl/

Indigenous languages of Latin America

The Archive of the Indigenous Languages of Latin America
https://ailla.utexas.org/

Romani

Russian Romani Corpus
http://web-corpora.net/RomaniCorpus/search/

Sign Languages

As part of the Language Archive (Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands)
https://www.mpi.nl/resources/data

Telephone conversations

Switchboard: a collection of about 2,400 two-sided telephone conversations among 543 speakers (302 male, 241 female) from all areas of the United States.
https://catalog.ldc.upenn.edu/LDC97S62

Translation

Multilingual Student Translation Corpus (MUST)
https://uclouvain.be/en/research-institutes/ilc/cecl/must.html

Evangelia Adamou, Paris (France)

53. Psycholinguistic methods in the study of bilingualism

1. Introduction
2. Measuring changes in attention to reveal the initial stages of language acquisition
3. Methods involving judgments: Two methods for examining sentence processing
4. Methods involving eye-movement recording
5. Methods involving response speed and accuracy measurements
6. Methods for studying simultaneous interpreting
7. Methods involving the manipulation of target processing by another stimulus
8. Methods involving (functional) neuroimaging
9. References

1. Introduction

In this chapter, we provide an overview of frequently used research methods in the psycholinguistic study of (individual) bilingualism. The study of individual bilingualism primarily examines the cognitive and neural mechanisms, processes, and knowledge structures involved in the acquisition and use of two (or more) languages. It also covers topics such as translation ability and the consequences of mastering two languages for other aspects of cognition (e.g. executive control and theory of mind).

The research methods employed in the study of bilingualism are largely similar to those used in more conventional psycholinguistics. Though by no means exhaustive − if only because researchers can be very creative in pursuing their goals − the collection of research methods presented in this chapter can be used to address questions concerning

https://doi.org/10.1515/9783110435351-053

the three main areas of study in bilingual psycholinguistics (i.e. language acquisition, comprehension, and production) and questions concerning several linguistic domains (e.g. the lexicon, syntax, and semantics).

Because most of the selected methods can be used to address several research questions, ordering of the methods based on research topics was not possible. Instead, we primarily ordered the methods according to the dependent variables and tasks involved, with each section title reflecting the dependent variable or task of interest. The section titles thus provide a way for scholars in the field to quickly find methods suitable for their purposes.

Space limits only allowed a succinct description of the presented methods. To enable readers to obtain further details about each method and to illustrate how it can be applied in bilingualism research, we have briefly described in each section a bilingual study that used the method in question. These brief descriptions of exemplary studies will likely provide readers with a clear sense of the content of the study of individual bilingualism as well as of the role(s) played by a set of major variables in this research area.

2. Measuring changes in attention to reveal the initial stages of language acquisition

2.1. Habituation-dishabituation paradigm

The habituation-dishabituation paradigm is a procedure to assess whether very young babies can discriminate between different stimuli, for example, between speech sounds from two languages. It is based on the well-known facts that babies and young children tend to scrutinize a new stimulus but lose interest when it is no longer new, for example, after it has been presented multiple times in a row. Babies' interest can be measured in various ways, for example, by measuring time spent looking at a stimulus (Weikum et al. 2007) or sucking rate on a pacifier (high amplitude sucking procedure; see Byers-Heinlein, Burns, and Werker 2010). Loss of interest, a reflection of habituation to a stimulus, is evident from a decrease in looking time or sucking rate. Once habituation is established, a new stimulus is presented to the baby. Only if the baby notices the difference will dishabituation occur, and as a result, looking times or sucking rates will increase. Using this method, the development of perception of speech sounds can be investigated with very young babies as participants (Christophe and Morton 1998). In a closely related paradigm, slightly older babies were rewarded for turning their heads upon detecting a difference between presented speech sounds (Werker and Tees 1999). Here, a head turn indicated that the babies could notice the difference between the sounds.

2.2. Preference procedures

Another way to assess language development in young children is the preferential looking procedure. This method does not rely on novelty detection, but rather establishes whether young children understand (parts of) the linguistic material they hear. Young

children are presented with two different visual stimuli (pictures, videos, or life materials) at two adjacent locations, and hear linguistic material that is congruent with one of these stimuli. If the child looks at the congruent-stimulus location more than at the other location, it can be inferred that the child has understood that the linguistic material referred to the congruent stimulus (Bobb et al. 2016).

In a related procedure, known as the disambiguation task, young children are given the choice to look at two screens presented next to each other, each displaying an object. The child knows the name of one of the objects, but the name of the other object is unfamiliar. When an (unknown) new object name is presented, and the child looks longer at the object with the unfamiliar name than at the one with the known name, researchers infer that the child is able to disambiguate the new word. By using this procedure, it has been shown that bilingual children aged 16–18 months exhibit different behaviors as compared to monolingual controls (Byers-Heinlein and Werker 2009).

3. Methods involving judgments: two methods for examining sentence processing

3.1. Thematic role assignment

Readers and listeners assign semantic roles to the words and phrases they read or hear. For example, upon reading the sentence *John kicked the dog*, the word *John* plays the actor (or agent) role, while the noun phrase *the dog* plays a recipient (or patient) role. One line of research using the thematic role-assignment method aims at investigating possible differences between monolingual speakers of different languages in preferred role assignments, and if such differences exist, between the particular role-assignment strategies bilinguals use. The question then becomes whether bilinguals always use the strategies associated with their L1, their L2, or a mixture thereof. In a study by Vaid and Chengappa (1988), for example, listeners heard three-word sentences, each consisting of two nouns and one verb, in various word orders. After hearing each sentence, they were asked to indicate which noun performed the action (reaction times were not measured). Under one condition, the words were taken from English, under the other, they were taken from Kannada (a language used primarily in the south of India). The participants could use noun animacy cues (i.e. whether the noun referred to an animate or inanimate entity), word order cues, and noun-verb agreement cues (in which one or both nouns agreed with the verb in number). The performance of a group of adult bilingual Kannada-English speakers was compared to that of a group of monolingual Kannada speakers. Under both the English and the Kannada condition, bilinguals showed a pattern of role assignment that was very similar to that of monolingual Kannada speakers (and very different from monolingual English speakers), indicating a transfer of role assignment strategies from Kannada to English.

3.2. Grammaticality judgments

When using this method, sentences are presented to participants and they are asked to indicate whether or not each sentence is grammatically correct in a particular language.

The method has been, for example, a useful tool in testing the idea that the learning of both the native and a foreign language might be different (and less successful) after a critical (or sensitive) period has ended rather than before such a period ends. In Birdsong and Molis (2001) learners of English heard sentences which were grammatically correct (e.g. *The horse jumped over the fence yesterday*) or incorrect (e.g. *The horse jumped the fence over yesterday*), and then circled a *yes* or *no* response about the grammatical correctness of each sentence on an answer sheet. The study replicated a study by Johnson and Newport (1989), who investigated speakers of L1 Korean and L1 Chinese learning English (whereas Birdsong and Molis [2001] tested a group of speakers of L1 Spanish learning English). The results showed, among others, that the total number of correct grammaticality judgments of participants who arrived in the United States after the closure of the putative critical (or sensitive) period correlated negatively with their age of arrival in the United States. This result, they argued, was not in line with the critical (or sensitive) period hypothesis.

Aside from response accuracy, response speed can also be measured using this task. Lago, Garcia, and Felser (2018) studied bilinguals who acquired German as an additional language. Participants read German sentences in which a pronoun and possessor either agreed or disagreed in gender, and indicated whether each sentence was acceptable or not. L1 Spanish-L2 English participants showed a smaller effect of pronoun-possessor agreement violation on response times than L1 English-L2 Spanish participants. In addition, although the overall accuracy of the judgments was the same for both groups, proficiency in English played a role within the L1 Spanish-L2 English group: participants with high proficiency in English performed better on this task than participants with low proficiency. These results were attributed to the fact that, unlike the English and German grammars, the Spanish grammar lacks pronoun-possessor agreement; still, knowledge of L2 English grammar increased sensitivity to pronoun-possessor agreement violations in the additional language (German).

4. Methods involving eye-movement recording

The recording of eye movements has proved a very useful technique in studying both visual and auditory language comprehension. In the visual domain, participants are presented with sentences or other text units, and recordings of eye movements reveal where and how long participants look while processing the presented text. From these measurements, the cognitive processes concerning word and sentence processing can be inferred (using the eye-mind assumption, which states that the mind processes the text fragment on which the eyes are currently fixated). The technique also allows researchers to study potential processing problems that cause participants to look back at previously read text, for example, because an initial interpretation of the meaning of a word or sentence proved to be wrong. For an example of the application of this technique in the study of bilingualism, see Frenck-Mestre and Pynte (1997), who studied how bilinguals resolved syntactic ambiguities.

In studying auditory language comprehension, eye-movement recordings have, for instance, been used to study whether words from both languages are simultaneously activated when bilinguals hear a word in one of their languages. In an application of the

technique known as the *visual world paradigm* (for a review, see Huettig, Rommers, and Meyer 2011) participants' eye movements are recorded while they see various pictures of objects and simultaneously hear a word. One of the pictures (the target picture) depicts the heard word. The crux of the method involves the manipulation of various aspects of the depicted non-target objects. For example, in Spivey and Marian (1999), one of the non-target pictures sometimes showed an object whose name in the bilingual participants' other language was phonetically related to the word spoken in the language of the experiment. Participants looked more at these non-target objects than at non-target objects with totally unrelated names, indicating that, during word recognition, words in both languages known by a bilingual are activated. Another recent example is a study by Morales et al. (2016), in which Italian-Spanish bilingual participants saw two pictures whose names always had congruent gender in Spanish, but either congruent or incongruent gender in Italian. Participants received Spanish-language auditory instructions that indicated a target object by using its article and its noun. Recordings of eye movement showed that participants looked less at target objects with incongruent Italian gender names than at target objects with congruent Italian gender names. The authors explained this effect as a result of the interaction of the bilingual's languages at the gender level.

5. Methods involving response speed and accuracy measurements

5.1. Lexical decision

On each trial in the lexical decision task, participants see a letter string and are asked to indicate whether it forms a word or not. Reaction times (RTs) and accuracy measures of the responses are the dependent variables. RTs are assumed to include the time required for lexical access, the component in which researchers are specifically interested. However, RTs also include the time involved in making (and executing) the decision. Many studies have shown that the decision-making process can be influenced by a diverse set of factors, including the percentage of *yes* responses and the composition of the letter strings that do not form a word.

In bilingualism research, participants can be instructed to produce a *yes* response if the letter string is a word in either of their two languages (*generalized* or *language-neutral* lexical decision), or in just one of them (*language-specific* lexical decision). Many studies using the lexical decision task have been devoted to examining the processing of *interlexical homographs* (an interlexical homograph is a written word that has different meanings in a bilingual's two languages) and *interlexical homophones* (an interlexical homophone is a spoken word that has different meanings in a bilingual's two languages). The results of such studies may help to answer questions regarding the activation of a language currently not in use in a bilingual lexicon, but the interpretation of the results is not always straightforward. For example, interlexical homographs may cause either longer, shorter, or similar reaction times compared to words that exist in the target language only, depending on the specifics of the task instruction (i.e. to perform a language-neutral or a language-specific lexical decision task) or the stimulus materials

used (Dijkstra, van Jaarsveld, and ten Brinke 1998). For a recent example of a study using this task, see Brysbaert, Lagrou, and Stevens (2017).

5.2. Identification

Methods involving identification most often use accuracy measures as the dependent variable. The method consists of the presentation of hard-to-identify word stimuli, and the task of the participant is to try to identify what was presented. Words may be made hard to identify by using visual masks and brief presentation of the words themselves, or by presenting only parts of them auditorily. As with the lexical decision task, identification tasks not only encompass processes devoted to the component of interest (lexical access), but also a decision process about what was perceived. In addition, visual or auditory processing itself may play an important role in these tasks.

In studying bilingual visual word recognition, Dijkstra, Grainger, and van Heuven (1999) used a *progressive demasking* technique in which they presented a masking stimulus with a presentation duration of 300 ms followed by a word stimulus with a presentation duration of 15 ms. Next, this presentation sequence was repeated but the mask presentation duration was decreased by 15 ms and the word presentation duration was increased by 15 ms. The adjustment of the durations was repeated until the participant indicated recognition of the stimulus word. Participants then typed the stimulus word at the computer. Mean RTs and error percentages were thus available for further analyses. The authors used English target words that could be related in meaning, orthography, or phonology to Dutch words. In comparison with unrelated control words, Dutch-English bilinguals were faster at recognizing the target words when they were related in meaning or orthography, but slower when they were related in phonology. The authors presented interpretations of these findings based on the structural characteristics of the bilingual processing system as well as the strategic use of this system.

Another identification technique is to embed a target word in a presentation sequence, in which the target is preceded by both a masking stimulus and a briefly presented prime stimulus (see Section 7.1), and followed by another masking stimulus. Using this technique, Brysbaert, van Dyck, and van de Poel (1999) presented French target words to Dutch-French bilinguals for 42 ms. The primes could be phonologically related to the target word or not (hence the name of this technique: *masked phonological priming*). Identification accuracy was the dependent variable. One of the results was that primes that were Dutch homophones of the French target words caused better recognition of those targets as compared to unrelated primes.

In an auditory word recognition study by Grosjean (1988), French-English bilingual participants heard French sentences whose presentation initially stopped at the onset of a target word. The target word was either an English guest word or a French borrowing thereof. In consecutive presentations of the sentence, more and more of the target word was revealed, in increments of 40 ms. This technique is known as *gating*. After each presentation, participants guessed the target word and indicated their confidence in the guess and their idea about the language of the target word. The results showed that phonotactic language cues and the composition of the French and English lexicons influenced the speed of word recognition.

5.3. Naming

A relatively straightforward task in the study of bilingualism is naming. Participants can be asked to read visually presented words aloud or to name pictures. In both tasks, RTs and/or accuracy measures can be recorded. Researchers use the reading aloud task primarily to measure lexical access time for the words involved. However, as with the lexical decision and identification tasks discussed above, it is well known that other variables might influence the results. For example, in languages with alphabetic scripts, participants might sometimes bypass lexical access and use grapheme-to-phoneme conversion rules to produce a response (as they do when reading non-words aloud). In addition, performance on this task also involves word production processes, the ease of which may vary independent of lexical access time. The latter processes may be identified by delaying the response until well after lexical access has taken place (a procedure known as *delayed naming*).

Using this task, de Groot et al. (2002) compared the influence of many linguistic variables (semantics, frequency, length, lexical neighborhood, word onsets, consonant clusters, sound intensity, and cognate status) in a study with Dutch-English participants. In addition to common word naming, a delayed word naming task and a lexical decision task were used; stimulus materials were derived either from Dutch or English. Onset structure turned out to be the most important variable governing naming times in Dutch, whereas frequency and length variables were particularly important in English. Comparisons of the naming and lexical decision tasks showed that, when these tasks were performed in Dutch, naming produced results that were quite different from those emerging in lexical decision (but quite similar to the results of delayed naming). When the tasks were performed in English, the results for naming and lexical decision were more similar. Overall, the authors concluded that the results of the naming and lexical decision tasks seem to reflect very different processes. In this particular experiment, words were presented in isolation, but they can also be presented in a sentence context (Schwartz and Kroll 2006). In the latter study, sentences were presented automatically word by word (by means of *rapid serial visual presentation*, or RSVP), but word presentation may also be controlled by the participant (*self-paced reading*).

Another naming task is picture naming. On each trial, participants see a picture on a screen and name the depicted object as quickly and accurately as possible. RTs are recorded by using a voice-key, and the experimenter judges whether or not the response was accurate. The picture naming task seems well suited, for example, to investigate whether words from a bilingual's not-in-use language are nevertheless activated when speaking in the other language. Costa, Caramazza, and Sebastián-Gallés (2000) investigated this issue by asking Spanish-Catalan bilingual participants to name pictures in Spanish. The names of the depicted objects shared phonology between the two languages (so-called *cognates*, e.g. Catalan *gat* and Spanish *gato*, English 'cat') or were unrelated (*noncognates*). The results were clear: pictures with cognate names were named more quickly than pictures with noncognate names, suggesting that Catalan words were nevertheless activated during naming in Spanish. In this study, the pictures were presented in isolation, but they can also be presented in a sentence context, as in Starreveld et al. (2014); this latter study employed a self-paced reading technique.

5.4. Word translation

One of the skills bilinguals possess is to translate between languages, that is, to transpose a language unit from one language (in translation studies, typically called the 'source' language) into another (the 'target' language). In a substantial number of translation studies, the language unit to translate is the individual word. Other translation research deals with translating larger linguistic segments in one form or another (e.g. from text to text; from speech to speech, as in simultaneous interpreting; see Section 6). Unlike the latter type of studies, word translation studies typically use RTs and accuracy measures as the dependent variables. One main goal of these studies has been to chart the factors affecting word translation (e.g. word characteristics such as concreteness and frequency; L2 proficiency) and explain the reasons why. Another goal is to reveal the architecture of the memory structures representing word knowledge in bilingual memory. In pursuing these goals, RTs obtained for translation from the stronger language (most often the L1) to the weaker L2 ('forward' translation) is often compared with RTs in 'backward' translation (from L2 to L1). In combination with a variable that can reveal whether word meaning is accessed in memory during the translation process, this translation-direction variable is thought to reveal the structure of lexical representations in bilingual memory, and the processing routes through these structures (i.e. circumventing or accessing meaning representations; Kroll and Stewart 1994).

5.5. Language switching

An important issue in bilingual research concerns bilinguals' ability to control which language to use, depending on the demands of the current communicative setting. One way researchers have examined this issue is by using language-switching methods. In most such studies, the switches are imposed by the experimenter. For example, Costa, Santesteban, and Ivanova (2006) used a picture naming task (see Section 5.3) in which bilingual participants named pictures in one or the other language, depending on the color in which the picture was displayed. Sometimes the color changed between successive trials (switch trials), while sometimes it did not (non-switch trials). Reaction times to switch and non-switch trials were then compared in order to assess the so-called switch costs (the RT difference between responses following a response in the same language versus one in the other language). The authors indeed found switch costs, but also showed that these costs were asymmetrical for unbalanced bilinguals: switching to the dominant native language resulted in greater switch costs than switching to the weaker second language. However, this asymmetry might be linked to the fact that switching was imposed by the experimenters, as the results stand in contrast to findings from Gollan and Ferreira (2009: 643), who instructed bilingual participants to "just say whichever name comes to mind most quickly" and obtained symmetrical switch costs for unbalanced and balanced bilinguals. Results such as these inform theories on language control in bilinguals.

6. Methods for studying simultaneous interpreting

A special case of language control concerns simultaneous interpreting, wherein most of the time, both languages must be used concurrently by the interpreter. Simultaneous interpreting encompasses multiple components, including source-language comprehension, target-language production, and memorizing portions of the input before they can be expressed in the target language. An oft-used approach to studying this complex task is to compare the performance of professional interpreters and 'ordinary' bilinguals on one or more of these components. For example, Christoffels, de Groot, and Kroll (2006) compared three groups of native Dutch speakers proficient in L2 English (trained interpreters, university students, and highly proficient English teachers) on two indices of working memory capacity (reading and speaking span) and two measures of word-retrieval efficiency (word-translation speed and picture-naming speed). They found that the interpreters had a larger working memory capacity than both the English teachers and the students, but were only faster than the students in word translation and picture naming in English. The authors concluded that fast word retrieval in an L2 emerges from extensive practice in the L2, and is thus concomitant with proficient bilingualism, but that a large working memory capacity results from specific on-the-job interpreting practice.

A common method for studying the various components of simultaneous interpreting in interaction with one another is by measuring the ear-voice-span (EVS; Goldman-Eisler 1972), the time between the start of a fragment in the source language and the moment the target-language equivalent of this fragment is rendered by the interpreter. The EVS is a sensitive measure of momentary processing load, with increases in processing demands lengthening the EVS. Other ways to assess interpreting performance are the use of various quality-of-performance measures, including the assessment of interpreting errors and durations of pauses.

7. Methods involving the manipulation of target processing by another stimulus

7.1. Priming of target processing by another stimulus

Priming is a method which can be combined with many other methods. The idea is simple: a target stimulus is preceded by another stimulus (the prime). Whenever responses to the target are influenced by prime presentation, processing of the target must have been facilitated by (or interfered with) as a consequence of processing of the prime.

Many variations on the procedure exist: the prime can be presented at various intervals before target presentation; the duration of the prime presentation can be varied; and the prime can be presented in full view or can be masked. In bilingual research, cross-language priming effects are often examined. In such research, the prime and the target are taken from a bilingual participant's different languages. We give a few examples. When a lexical decision task (see Section 5.1) is performed on the target and if the prime and the target share a semantic relationship, cross-language semantic priming can be obtained (Frenck and Pynte 1987); when the prime and the target are translation equivalents, translation (or cross-language repetition) priming effects can be obtained (Altarriba

1992). When a perceptual identification task (see Section 5.2) is performed on the target and if the prime and the target share a phonological relationship, cross-language phonological priming can be obtained (Brysbaert, van Dyck, and van de Poel 1999). Such effects indicate that language representations and/or processing operations are shared between bilinguals' two languages.

7.2. Interference of target processing by another stimulus

An important method that has been used to investigate (bilingual) language production is to deliberately induce interaction between production and perception processes. Both processes are important in the picture-word interference task. In this task, the main component is picture naming (see Section 5.3), but now a picture is accompanied by a distractor (often a word). Participants are asked to name the pictures and to ignore the distractors. Despite this instruction, effects of various relationships between the distractor and the target (i.e. the depicted object's name) can be obtained. Again, several variants of the task exist. For example, the distractors can be presented visually or auditorily. Often, the distractors are presented slightly before, concurrently with, or slightly after picture presentation. In this way, a time course of the effect under investigation can be obtained. An example of the use of this method is presented in Costa, Miozzo, and Caramazza (1999). Catalan-Spanish bilinguals named pictures in Catalan. Both Spanish and Catalan distractors were used, which could be semantically related to the target, unrelated to it, or its translation equivalent. The authors found, among other results, cross-language interference effects of semantically related distractors, as well as cross-language facilitation effects of translation equivalents (both effects were relative to a baseline obtained using unrelated distractors). They explained their results by assuming that representations of words from both languages were activated during task performance, but that lexical selection was restricted to representations of words from the response language only.

Instead of a target picture, a to-be-named or to-be-translated target word can also be used (see Sections 5.3. and 5.4). The difference between target and distractor may then be indicated by a difference in color, position, or time of presentation.

8. Methods involving (functional) neuroimaging

Functional neuroimaging methods are different from all methods mentioned above because they do not measure overt behavior, but rather brain activity that occurs during stimulus processing. Though this is the quintessential feature of these methods, one should always bear in mind that brain activity and stimulus processing are not necessarily causally related. In addition, although heightened activity in a given brain region is often interpreted as showing evidence of stimulus processing by that region, absence of increased brain activity in a given region does not necessarily mean the stimulus is not processed by that region, especially when stimulus processing is (largely) automatic. Still, the methods described in this section have proved very useful in many research areas, including bilingual studies. Although other neuroimaging methods exist (e.g. mag-

netoencephalography, or MEG; near-infrared spectroscopy, or NIRS; positron emission tomography, or PET), we discuss below only the more frequently used methods.

8.1. Using electrical signals

Using electrodes that are placed at the scalp, it is possible to record electrical brain activity directly and continuously. Neural communication in the brain involves the transfer of ions, and the movement of large amounts of ions can be detected as they cause very slight changes in voltage. The detection process, and the graphical representation thereof, is called electroencephalography (EEG). A derivative of the EEG signal (not the signal itself), the *event-related potential* (ERP), is particularly informative to researchers interested in language processing (and other aspects of cognition). This derivative involves combining and averaging the EEG signals following, and time locked to, the repeated presentation of a certain type of stimulus event.

For language scholars, the most important effects (i.e. peaks in the ERP signal, typically called 'components') that have emerged from this field are the N400 and the P600. The N400 is a negative peak in the ERP waveform that occurs 400 ms after a semantic abnormality in the stimulus is presented (Kutas and Hillyard 1980; for a broader interpretation of the N400 effect, see Lau, Phillips, and Poeppel 2008). The P600 is a positive peak in the ERP signal that occurs 600 ms after the presentation of a syntactic abnormality in the stimulus (Osterhout and Holcomb 1992). Both peaks are relative to the averaged waves produced by stimuli that do not show semantic or syntactic abnormalities. The temporal resolution of the EEG method is high, whereas its spatial resolution is rather low.

Ardal et al. (1990) and Weber-Fox and Neville (1996) were among the first to adopt the EEG/ERP methodology in bilingual studies. Ardal and colleagues (1990) compared the N400 in English monolinguals and English-French bilinguals. Participants saw sentences which were semantically normal (e.g. *It is raining and I forgot my umbrella*) and anomalous (e.g. *I generally like menthol bottles*). The words of each sentence were presented one at a time, and ERP measurement started upon presentation of the final word, which was either semantically normal or anomalous. The authors found that the N400 occurred earlier in monolingual than in bilingual participants, especially for bilinguals tested in their second language. They attributed these results to a higher degree of automaticity of language processing in monolinguals than in bilinguals, and in bilinguals' processing of their L1 as compared to their L2. In addition to semantic anomalies, Weber-Fox and Neville's (1996) study also included syntactic anomalies. The participants were presented with grammatically incorrect sentences (e.g. *The scientist criticized Max's of proof the theorem*). Monolingual English speakers showed a P600 peak following the grammatical violation. For Chinese late learners of English, the amplitude of the P600 peak was reduced, suggesting that such learners either did not attempt to repair the meaning of the sentence, or did so later. For a review of ERP research in bilingual language processing, see Moreno, Rodríguez-Fornells, and Laine (2008); for a recent example, see Martin et al. (2016).

8.2. Using hemodynamic magnetic signals

Another way to measure brain activity is to register the amount of blood flow in particular areas of the brain. The technique exploits the fact that increased neuronal activity in the brain is followed by an increased oxygen-rich blood flow beginning around 2 s later and peaking around 4–6 s later, after which blood flow falls back to – and typically undershoots – the original level. In functional magnetic resonance imaging (fMRI), blood oxygen level dependent (BOLD) contrast is measured by studying changes in the magnetic fields induced by the MRI scanner. The BOLD contrast is rooted in the fact that oxygenated blood is not magnetic, whereas deoxygenated blood is. When a change in the oxygen concentration of blood flow occurs, a difference in the resultant magnetic fields can be detected, which researchers then attribute to neural activity at a specific location in the brain. Researchers usually measure brain activity for a time duration of around 20 to 30 s, during which participants perform a certain task. In a variation on this method, measuring of the BOLD response is time-locked to the presentation of a single stimulus, and the results of several measurements are then averaged. This variant is known as event-related fMRI. In general, in comparison with ERP, fMRI has low temporal resolution (because the blood flow needs time to build up), but high spatial resolution: brain structures of a few mm^3 can be differentiated.

One of the issues addressed in bilingual research using fMRI is whether the two languages of bilinguals are processed by the same or different brain areas. In such studies, participants perform various language tasks (e.g. word-production, covert sentence generation, covert word-stem completion, semantic categorization, listening, or silent reading). Comparisons can then be made between imaging results for various tasks used in the study, often including a condition in which participants do not actively process linguistic material (e.g. the participant looks at a fixation stimulus). The results of such studies diverge, but a review of the literature by Indefrey (2006) showed that, in general, bilinguals' L1 and L2 are not processed by different brain areas. However, a number of the areas involved (15 out of a total of 114, that together cover the entire brain) did show differences in the level of activation between L1 and L2, with L2 processing causing greater activation than L1 processing in most cases (13 out of 15 areas).

Another research area in the study of bilingualism concerns the control processes that allow a bilingual to selectively use one language, or to switch between languages if needed. Many such studies indicate that the dorsolateral prefrontal cortex, the anterior cingulate cortex, and the caudate are involved in language control, in both production (Abutalebi et al. 2007) and comprehension (van Heuven et al. 2008) tasks.

8.3. Diffusion tensor imaging

Finally, we describe an example of a neuroimaging method that allows mapping of (changes in) structural connectivity between language areas of the brain. Water molecules diffuse more easily parallel to a white matter tract than perpendicular to it. The direction of diffusion of water molecules can be measured with a technique derived from magnetic resonance imaging called diffusion tensor imaging (DTI). From DTI measurements, the precise location and mass of white matter in the brain can be reconstructed. Xiang et al. (2015) used DTI to investigate the changes in white matter in language-

related areas of the brain during acquisition of a second language. German participants took an intensive 6-week Dutch language learning course, and possible pathways between the left parietal lobe and four subregions of Broca's complex were mapped both before and after the course, and in both the left and the right hemisphere. Dutch proficiency was measured with a cloze test. The results showed that during second language learning, lateralization dominance of a pathway mainly along the arcuate fasciculus shifted from the left to the right, and back again, with increasing proficiency.

9. References

Abutalebi, Jubin, Jean-Marie Annoni, Ivan Zimine, Alan J. Pegna, Mohamed L. Seghier, Hannelore Lee-Jahnke, François Lazeyras, Stefano F. Cappa & Asaid Khateb
2007 Language control and lexical competition in bilinguals: An event-related fMRI study. *Cerebral Cortex* 18(7). 1496–1505.

Altarriba, Jeanette
1992 The representation of translation equivalents in bilingual memory. In Richard J. Harris (ed.), *Cognitive processing in bilinguals*, 157–174. New York: Elsevier Science.

Ardal, Sten, Merlin W. Donald, Renata Meuter, Shannon Muldrew & Moira Luce
1990 Brain responses to semantic incongruity in bilinguals. *Brain and Language* 39(2). 187–205.

Birdsong, David & Michelle Molis
2001 On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language* 44(2). 235–249.

Bobb, Susan. C., Laila Y. Drummond Nauck, Nicole Altvater-Mackensen, Katie Von Holzen & Nivedita Mani
2016 Listening with your cohort: Do bilingual toddlers co-activate cohorts from both languages when hearing words in one language alone? In John W. Schwieter (ed.), *Cognitive control and consequences of multilingualism*, 47–70. Amsterdam: John Benjamins.

Brysbaert, Marc, Evelyne Lagrou & Michaël Stevens
2017 Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Bilingualism: Language and Cognition* 20(3). 530–548.

Brysbaert, Marc, Goedele van Dyck & Marijke van de Poel
1999 Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance* 25(1). 137–148.

Byers-Heinlein, Krista & Janet F. Werker
2009 Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science* 12(5). 815–823.

Byers-Heinlein, Krista, Tracey C. Burns & Janet F. Werker
2010 The roots of bilingualism in newborns. *Psychological Science* 21(3). 343–348.

Christoffels, Ingrid K., Annette M. B. de Groot & Judith F. Kroll
2006 Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language* 54(3). 324–345.

Christophe, Anne & John Morton
1998 Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science* 1(2). 215–219.

Costa, Albert, Alfonso Caramazza & Nuria Sebastian-Galles
2000 The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(5). 1283–1296.

Costa, Albert, Michele Miozzo & Alfonso Caramazza
1999 Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language* 41(3). 365–397.
Costa, Albert, Mikel Santesteban & Iva Ivanova
2006 How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 32(5). 1057–1074.
de Groot, Annette M. B., Susanne Borgwaldt, Mieke Bos & Ellen van den Eijnden
2002 Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language* 47(1). 91–124.
Dijkstra, Ton, Henk van Jaarsveld & Sjoerd ten Brinke
1998 Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language & Cognition* 1(1). 51–66.
Dijkstra, Ton, Jonathan Grainger & Walter J. B. van Heuven
1999 Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory & Language* 41(4). 496–518.
Frenck, Cheryl & Joel Pynte
1987 Semantic representation and surface forms: A look at across-language priming in bilinguals. *Journal of Psycholinguistic Research* 16(4). 383–396.
Frenck-Mestre, Cheryl & Joel Pynte
1997 Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology* 50A. 119–148.
Goldman-Eisler, Frieda
1972 Segmentation of input in simultaneous translation. *Journal of Psycholinguistic Research* 1(2). 127–140.
Gollan, Tamar H. & Victor S. Ferreira
2009 Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 35(3). 640–665.
Grosjean, François
1988 Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes* 3(3). 233–274.
Huettig, Falk, Joost Rommers & Antje S. Meyer
2011 Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica* 137(2). 151–171.
Indefrey, Peter
2006 A meta-analysis of hemodynamic studies on first and second language processing: Which suggested differences can we trust and what do they mean? In Marianne Gullberg & Peter Indefrey (eds.), *The cognitive neuroscience of second language acquisition*, 279–304. Blackwell Publishing, Malden, MA.
Johnson, Jacqueline S. & Elissa L. Newport
1989 Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21(1). 60–99.
Kroll, Judith F. & Erika Stewart
1994 Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33(2*)*. 149–174.
Kutas, Marta & Steven A. Hillyard
1980 Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207(4427). 203–208.

Lago, Sol, Anna Stutter Garcia & Claudia Felser
2018 The role of native and non-native grammars in the comprehension of possessive pronouns. *Second Language Research.* https://doi.org/10.1177/0267658318770491. (accessed on 30 January 2019).
Lau, Ellen F., Colin Phillips & David Poeppel
2008 A cortical network for semantics: (de) constructing the N400. *Nature Reviews Neuroscience* 9(12). 920–933.
Martin, Clara D., Xavier Garcia, Audrey Breton, Guillaume Thierry & Albert Costa
2016 World knowledge integration during second language comprehension. *Language, Cognition and Neuroscience* 31(2). 206–216.
Morales, Luis, Daniela Paolieri, Paola E. Dussias, Jorge R. Valdés Kroff, Chip Gerfen & María Teresa Bajo
2016 The gender congruency effect during bilingual spoken-word recognition. *Bilingualism: Language and Cognition* 19(2). 294–310.
Moreno, Eva M., Antoni Rodríguez-Fornells & Matti Laine
2008 Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics* 21(6). 477–508.
Osterhout, Lee & Phillip J. Holcomb
1992 Event-related brain potentials elicited by syntactic anomaly. *Language and Cognitive Processes* 31(6). 785–806.
Schwartz, Ana I. & Judith F. Kroll
2006 Bilingual lexical activation in sentence context. *Journal of Memory and Language* 55(2). 197–212.
Spivey, Michael J. & Viorica Marian
1999 Crosstalk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science* 10(3). 281–284.
Starreveld, Peter. A., Annette M. B. de Groot, Bart M. M. Rossmark & Janet G. van Hell
2014 Parallel language activation during word processing in bilinguals: Evidence from word production in sentence context. *Bilingualism: Language and Cognition* 17(2). 258–276.
Vaid, Jyotsna & Shyamala Chengappa
1988 Assigning linguistic roles: Sentence interpretation in normal and aphasic Kannada-English bilinguals. *Journal of Neurolinguistics* 3(2). 161–183.
van Heuven, Walter J. B., Herbert Schriefers, Ton Dijkstra & Peter Hagoort
2008 Language conflict in the bilingual brain. *Cerebral Cortex* 18(11). 2706–2716.
Weber-Fox, Christine M. & Helen J. Neville
1996 Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience* 8(3). 231–256.
Weikum, Whitney M., Athena Vouloumanos, Jordi Navarra, Salvador Soto-Faraco, Núria Sebastián-Gallés & Janet F. Werker
2007 Visual language discrimination in infancy. *Science* 316(5828). 1159.
Werker, Janet F. & Richard C. Tees
1999 Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology* 50. 509–535.
Xiang, Huadong, Tessa Marije van Leeuwen, Dan Dediu, Leah Roberts, David G. Norris & Peter Hagoort
2015 L2-proficiency-dependent laterality shift in structural connectivity of brain language pathways. *Brain Connectivity* 5(6). 349–361.

Peter A. Starreveld, Amsterdam (The Netherlands)
Annette M. B. de Groot, Amsterdam (The Netherlands)

54. Research ethics in contact linguistics

1. Introduction
2. Legal and regulatory aspects of ethics
3. Some moral aspects of ethics
4. Practical aspects of ethics
5. Interim summary
6. Ethics and research in contact linguistics
7. Summary
8. References

1. Introduction

Rumsey (2010: 105) defines ethics simply: "ethics necessarily involves acting with regard for another person or people". Bowern (2015: 167), writing on ethics in linguistic fieldwork, presents a general definition: "a way of working that you, the research community, and the language community think is appropriate". In a broad sense, ethics involves acting with regard for another person or people, as Rumsey puts it, but it also involves acting with regard for languages and for knowledge systems, with implications for who does research and for their research methods.

In this chapter, I focus on questions of ethics to be asked in undertaking research on language contact. Contact is the focus of attention of several groups − historical linguists, formal linguists, functional linguists, sociolinguists, psycholinguists, linguistic anthropologists − each with their own questions. Contact linguistics draws on various methods, including the comparative-historical method, grammatical analysis, sociolinguistic methods, and ethnographic methods such as participant observation, interviews, and conversations. One important way of obtaining data that is of value in contact linguistics is through fieldwork, an area where ethics is front and center, and which is my focus in this chapter.

Ethics is of particular concern in contact linguistics fieldwork since, as Garrett (2004: 48−49) points out, encounters under situations of contact may be "relatively egalitarian or marked by significant asymmetries of power […]; peaceful and mutually beneficial or coercive, exploitative, and otherwise detrimental to one or more of the groups involved [… and] situated and contingent, and hence fundamentally indeterminate". In other words, the social, political, and linguistic ecology of individuals and their communities matters (acknowledging the notorious complexity of the term 'community', and not further addressing the problems therewith here). In working with speakers in contact situations, ethical principles are critically important, and these are the topic of this chapter.

I focus on how an understanding of ethics has developed in fieldwork broadly, drawing primarily on sociolinguistic, ethnographic, and language documentation literature. I begin with an overview of three aspects of ethics: legal and regulatory, moral, and practical. I then turn to some aspects of ethics that are of particular importance in contact situations.

2. Legal and regulatory aspects of ethics

Modern attention to research ethics results from the Nuremberg Code (1947), developed out of the Nuremberg trials following World War II. This code, judged by Shuster (1997:

https://doi.org/10.1515/9783110435351-054

1436) to be "the most important document in the history of ethics of medical research", introduces a set of ethics principles for medical experimentation with human subjects. The code included principles such as informed consent, positive results for society, and avoidance of harm, seeking to balance experimentation with protection. The Declaration of Helsinki (WMA 1964–2013), first adopted in 1964 with several revisions since then, is another important document on human research ethics. The *Belmont Report: Ethical principles and guidelines for the protection of human subjects of research* (National Commission 1979) identifies as core principles respect for persons, beneficence, and justice, as well as the importance of informed consent, assessment of risks and benefits, and selection of subjects in the application of these principles.

Current ethics protocols have grown out of such documents. They began as principles for medical research, and have been extended to other areas where research involves human participants, including linguistic research. The principles are framed differently in different places; for example, the code regulating research with humans in Canada (Secretariat 2014) takes respect for human dignity as its underlying value. This is expressed through three core principles: respect for persons, concern for welfare, and justice.

In many places (countries, states, provinces, and/or communities), regulatory codes exist for research with human participants. In general, these aim for justice, respect for persons, and beneficence, expressed as questions about voluntary and informed consent, conflicts of interest, research methods, participants and their vulnerability, recruitment, compensation, risks and benefits, withdrawal, confidentiality, privacy, and data security, access, and retention.

Informed and voluntary consent, the cornerstone of all ethics frameworks, can be particularly complex. In fieldwork, consent tends to be an ongoing process rather than a one-time event. Do the researcher and the participants in the research have the same understanding of what the research is about, what it is not about, and what its potential consequences are? Is the researcher honest about what the research is truly about, or are the goals masked and presented in a way that allows the researcher to obtain the most natural linguistic data possible without being explicit that natural linguistic data is their goal? The role of informed consent in fieldwork is sometimes muddied. In language documentation, participants are generally informed about the major purpose of the fieldwork. However, in sociolinguistic research, the Observer's Paradox ("To obtain the data that is most important for linguistic theory, we have to observe how people speak when they are not being observed" [Labov 1972: 113]) intersects with informed consent. For instance, Eckert (2013: 15) notes that in her early fieldwork, she did not tell people she was interested in language rather than ethnography, and she asks at several years remove whether this was appropriate. Critiques of the Observer's Paradox are summarized in Schilling (2013: 115–116, 128), and Trechter (2013: 33) discusses what she identifies as "methodological machinations" in attempting to record authentic vernacular speech, including surreptitious recording and requests for emotional stories.

As many have written (e.g. Bowern 2010a; Eckert 2013), the regulatory framework can be both a help and a hindrance in fostering ethical research, and is often an uneasy fit with fieldwork. It is nevertheless important to be aware of the regulatory aspects of research ethics, whether the regulations come from governments, communities, or granting councils. However, at least as important is understanding the moral principles that underlie the regulatory principles, and I examine some of these in the next section.

3. Some moral aspects of ethics

I open with two moral perspectives on ethics: ethics with respect to languages, and ethics with respect to knowledge sharing and ownership. I then turn to models of ethical research, and finally, to language ideologies.

3.1. Ethics and languages

With the realization in the 1990s that many languages were not being passed on to the younger generation, there have been calls for the responsibility of the profession to study languages that are in danger of not being transmitted. Such situations necessarily involve contact, but the focus until recently has generally been largely on documenting languages before they cease to be transmitted, rather than on language contact. The drive to study languages is stated forcefully in an important article introducing language endangerment by Krauss (1992: 8):

> Seeing the present situation, I think that, at the very least, it behooves us as scientists and as human beings to work responsibly both for the future of our science and for the future of our languages, […] for the sake of posterity. What we need to do now stares us in the face. […] We must obtain adequate information on the condition of the languages of the world, better than we have now, and use it to plan priorities for linguistic work in a rational and coordinated way. […] Obviously, for scientific purposes, it is most urgent to document languages before they disappear.

Krauss focused on endangered languages, situations that almost by definition involve language contact and contact languages. While Krauss' statement has been highly influential, the ethics of the documentation of endangered languages has been critiqued. The goal of documentation at least began as one of documentation of "fluent monolingual speech, preferably employing 'classical' grammatical features and talking about 'traditional' topics" (Moore, Pietikäinen, and Blommaert 2010: 16). This is in conflict with another goal of documentation: recording examples of "how the people actually communicate with each other" (Himmelmann 2006: 7). (See Dobrin and Berson [2011: 194] for discussion.) What is ethical? Is it the ethical responsibility to record the language as it was spoken to the degree possible? Or is it to record the language in the way it is actually used? Research on a language as such and its structure might favor the former; research on language use and choice might favor the latter.

3.2. Ethics, knowledge sharing, and ownership

I now consider ethics with respect to knowledge sharing and ownership. Labov (1982) formulates principles that are well known in sociolinguistic research, and that speak to the importance of social action with respect to knowledge. One of them, the Principle of Error Correction, states an ethical obligation to ensure that knowledge gained through research is made publicly available.

> Principle of Error Correction
>
> A scientist who becomes aware of a widespread idea or social practice with important consequences that is invalidated by his own data is obligated to bring this error to the attention of the widest possible audience. (Labov 1982: 172)

The second of Labov's principles, the Principle of Debt Incurred, complements the first one, and focuses on the community from which the knowledge is obtained.

> Principle of Debt Incurred
>
> An investigator who has obtained linguistic data from members of a speech community has an obligation to use the knowledge based on that data for the benefit of the community, when it has need of it. (Labov 1982:173)

Wolfram (1993: 227) points out that these two principles represent *reactive* advocacy, where a linguist who has conducted research in a community "is called upon or feels obligated to respond to some erroneous sociolinguistic assumption or conclusion about the language of the community as the language is threatened socio-politically or socio-educationally". He proposes an additional principle, a *proactive* one, the Principle of Linguistic Gratuity.

> Principle of Linguistic Gratuity
>
> Investigators who have obtained linguistic data from members of a speech community should actively pursue positive ways in which they can return linguistic favors to the community. (Wolfram 1993: 227)

This principle is, as Wolfram suggests, an extension of those proposed by Labov. Wolfram (1993: 227–228) notes that linguistics researchers should "creatively search for a community-based collaborative model to return linguistic favors".

Underlying the three principles presented above is an assumption about the complex issue of ownership of knowledge, or intellectual property rights. While informed consent was probably given, Labov's principles assume that ownership of data rests with the researcher, to be used for public education and benefit, as determined by the researcher. It is less clear with Wolfram's (1993) Principle of Linguistic Gratuity where ownership lies: with the researcher, the participants, or both.

Ownership is one of the thorny issues that underlies much work on ethics in fieldwork. Eckert (2013: 13) succinctly highlights three critical questions with respect to ownership in fieldwork-based research:

1. The researcher's ownership and control of research ideas, hypotheses, procedures, results, and interpretation.
2. The extent to which the researched are disinterested subjects of the researcher's activity.
3. The responsibility of the researcher to seek and use research knowledge for the benefit of the researched.

These are important questions to keep in mind as one thinks about field-based research. The question of ownership becomes increasingly critical in a world where there is pressure for research data to be made openly available. Seeking to establish a truly shared

understanding of what this means is vital, as there are social and political reasons why data collected in the study of language contact might potentially bring risk and/or harm to the individuals involved.

3.3. Models of research

Questions surrounding topics such as ownership have played out in different ways. Cameron et al. (1992) identify three types of research models: ethical research, advocacy research, and empowerment research. At the core of these models are notions of power, ownership, and control.

Ethical research involves minimizing damage and inconvenience to research participants, while also acknowledging their contributions. Ethical research is often characterized with the preposition 'on'. In the words of Cameron et al. (1992: 15): "[i]n ethical research […] there is a wholly proper concern to minimize damage and offset inconvenience to the researched, and to acknowledge their contribution […] But the underlying model is one of 'research *on*' social subjects. Human subjects deserve special ethical consideration, but they no more set the researcher's agenda than the bottle of sulphuric acid sets the chemist's agenda". Ethical research might invoke Labov's Principle of Error Correction, introduced earlier. This is a type of ethical model focused on knowledge, and a perspective focused on the researcher and their responsibilities to collect and analyze data while treating research participants well. The control ultimately rests with the researcher.

Advocacy research adds a new component, with the researcher using their skills to advocate for research participants. According to Cameron et al. (1992: 15),

> The 'advocacy' position is characterized by a commitment on the part of the researcher not just to do research *on* subjects but research *on* and *for* subjects. Such a commitment formalizes what is actually a rather common development in field situations, where a researcher is asked to use her skills or her authority as an 'expert' to defend subjects' interests, getting involved in their campaign for healthcare or education, cultural autonomy or political and land rights, and speaking on their behalf.

Labov's (1982: 173) Principle of Debt Incurred is similar in force, suggesting the obligations of the researcher to the people with whom they work. While designed to be proactive, Wolfram's Principle of Linguistic Gratuity is an extension, with the researcher being seen as responsible to the community for using the knowledge gained for the benefit of the community by returning it to the community. This can be in the form of recording stories, making videos, and so on (e.g. Wolfram 2013). Writing on social ethics for sociolinguistics, Trechter (2013: 40) characterizes such a model as still involving inequality between the researcher and participants: it remains researcher controlled, with the researcher determining communities' needs.

Empowering research is understood by Cameron et al. (1992: 15) as research 'on', 'for', and 'with'. One of the things they "take that additional 'with' to imply is the use of interactive or dialogic research methods, as opposed to the distancing or objectifying strategies positivists are constrained to use. It is the centrality of interaction 'with' the researched that enables research to be empowering in our sense". The addition here of

'with' is important, and this model has been particularly favored by those doing fieldwork in endangered language situations. This model finds a more central place for community members. Nevertheless, as Czaykowska-Higgins (2009) points out, this model remains researcher-focused; in Trechter's terms, it would likely be characterized as involving inequality.

Czaykowska-Higgins (2009: 24) adds a fourth model, what she calls *Community-Based Language Research* (CBLR) and describes as "[r]esearch that is *on* a language, and that is conducted *for*, *with*, and *by* the language-speaking community within which the research takes place and which it affects. This kind of research involves a collaborative relationship, a partnership, between researchers and (members of) the community within which the research takes place". As Czaykowska-Higgins (2009: 25) notes, this model involves the researcher training community members to do the research, as well as the researcher being trained by (and learning from) community members on "issues related to language, linguistics, and culture, as well as about how to conduct research themselves appropriately within the community". Czaykowska-Higgins (2009: 25) continues:

> Thus, CBLR is based on the recognition that community members have expertise and can be experts. Because it recognizes that linguists are neither the sole researchers nor the only experts and that their role is to be partners in a collaborative relationship in which all partners learn from each other, the Community-Based Language Research model goes further than the Advocacy and Empowering research models in breaking down the boundary between researchers and language-users and/or community members.

Such a model might be characterized by Trechter as involving equality. See also e.g. Bischoff and Jany (2018), Chelliah and Reuse (2011), Dwyer (2006), Holton (2009), Leonard and Haynes (2010), Matras (2005), Rice (2006, 2009, 2010, 2011), and references within those works, for additional perspectives on these issues.

While recognizing that they are abstractions, thinking in terms of models is useful. Which model is appropriate will vary over space and time. One important consideration in current thinking is that a research model should not be assumed by the researcher, but rather negotiated with a community. With regard to the questions with which this section opened, the models generally move away from one where the researcher controls the ideas, hypotheses, procedures, and interpretations, to one where the participants become more true participants in the research, with the research serving their perceived needs. There is a shift from a narrow focus on knowledge, to the distributed responsibilities of all participants and ownership of knowledge (and possibly with restrictions on what the researcher is allowed to share).

The development of empowering, community-based models has been especially strong in research on endangered languages, as noted above, where issues of control and power are particularly evident. It is important to keep in mind that there is no 'one-size-fits-all' model of research. According to Trechter (2013: 33) "[…] consideration of ethics in linguistic research is necessarily framed within a language community and consistently foregrounds issues of shared rights, obligations, and responsibilities". Besnier (2013: 48) summarizes this as follows:

> While no simple ethical guideline applies to all situations, one can suggest a rule of thumb: prior to beginning a study, researchers must think through potential dynamics and ethical

> conflicts that may arise in the field, and all decisions about the nature, form, and circulation of research must be made with recognition of the complexities of the situation and the instability of power relations. […] In short, all sociolinguists must consider carefully questions of ethics broadly defined, taking into account not only issues of consent, but also power, scale, representation, subjectivity, and positionality.

Questions of power and control are prominent in much linguistic research, and consideration of issues of social justice and relationality have come to be viewed as critical to such research. To repeat, there is no 'one-size-fits-all' model, as what is considered 'socially just' can vary over space and time. The variety of potential research models suggests that one ethical responsibility of a researcher is to try to understand what the community in which they are working expects. This might occur through the regulatory process, but is more likely to happen through consideration of questions about what it means to work on a topic like language contact that so critically involves social issues, in addition to taking an ethnographic approach to the research, listening, and trying to hear what is actually being said (e.g. Dobrin and Berson 2011; Dobrin and Schwartz 2016; Gal 1979; Jaffe 1999; Kroskrity 2004; Kulick 1992, 1998).

3.4. Ethics and knowledge systems: language ideologies

Much has been written on language ideologies, or "beliefs, or feelings, about languages as used in their social worlds" (Kroskrity 2004: 498). Irvine and Gal (2000: 36), drawing on examples from language contact situations, write that there is "no 'view from nowhere,' no gaze that is not positioned", signaling the complexities of studying language contact and understanding the diverse ideologies that underlie such study, as well as the set of often contradictory ideologies and implicit assumptions that both researcher and participants bring to the research. See, for instance, Kroskrity (2004) and Sallabank (2013) for recent overviews of the literature on language ideologies, and the challenges that ideologies can bring with regard to knowledge.

Grenoble and Whitecloud (2014) address differences in ideologies regarding methods for obtaining knowledge between Western linguists and Indigenous groups. They argue that the Western approach favors extraction of language from its context, while the Indigenous approach is more holistic in nature. Grenoble and Whitecloud make clear how easy it is for Western researchers to impose their own Western conceptions of science as well as of local knowledge and beliefs without hearing what local participants are actually saying. One aspect of knowledge systems that is characteristic of many contact situations involves different recognitions of what it means to 'speak a language', with different accompanying interpretations of this, both by external researchers and by different members of a given community.

4. Practical aspects of ethics

The practical aspects of ethics take many forms. They may begin with the ethics protocols required by many universities and communities, responding to questions about top-

ics such as: recruitment, informed consent, sensitivity of research material, potential risks and benefits, deception, confidentiality, anonymity, and so on, as detailed in Section 2.

Other practical questions might arise in the course of fieldwork. Some of these might not seem like questions of ethics, but the consequences of *not* thinking about such issues could well have ethical consequences. Here, I summarize what is presented in Rice (2012) as direct questions for which there are not necessarily simple answers, with much depending on time, place, relationships, and other factors; see Macaulay ([2004] 2012) for a detailed, personal discussion of the realities of fieldwork, and what students should know before they undertake fieldwork:

> Who do I work with?
> What does the community that I am working with expect of me?
> What do I tell people about my funding?
> How do I tell people what I am doing?
> Should I pay the people who work with me?
> How much should I pay?
> How do I get informed consent?
> What can I record?
> Can I insist on video recording?
> What can I do with the data?
> While best academic practice today involves archiving data and making it freely available, is that possible, given community circumstances?

Other issues that may have ethical consequences might seem trivial at first glance. These include questions such as the following:

> Where should I live?
> What should I eat?
> How should I dress?

Furthermore, issues of safety of researchers and participants both could be involved, and if inappropriate decisions are made, unexpected social and/or political consequences. Balancing individual needs with community expectations can be a challenge, but it generally turns out to be valuable to have given thought to what might seem like 'small things', as pointed out in Rice (2012: 416).

5. Interim summary

Stressing the importance of attending to social relationships, Trechter (2013: 44) writes that "[b]ecause sociolinguistic ethics are most appropriately grounded in the social, the roles of all speech participants, as well as their institutions, culture, and history, play a vital role in determining the best methods for thoughtful, sustained interaction". Besnier (2013: 48) underscores this point, noting that "while researchers have the power to animate certain voices but not others, and to privilege certain representations over others, social groups and individuals also have the power to accept, resist, and reject representations of their social practices, particularly in a world in which the boundary between researcher, object of research, and audience is no longer straightforward". He and many

others stress the importance of considering issues of consent, power, scale, representation, subjectivity, and positionality. In short, there is no simple, singular recipe for 'ethical research': it involves ongoing negotiations and personal interrogation of respect for relationships, methods, and ways of knowing.

6. Ethics and research in contact linguistics

The social situations that are important to understand in studying contact linguistics are human situations, and thus subject to the ethical challenges of the social sciences more broadly, and of sociolinguistics and fieldwork in particular: i.e. the challenges of informed consent, clarity as to research purpose, doing some good, and confidentiality (see e.g. Schilling [2013] for discussion of ethical issues in sociolinguistic fieldwork, and Bowern [2015] and Chelliah and Reuse [2011] on fieldwork more broadly).

Are there things that make language contact situations special with respect to ethics? It is helpful to consider this question in light of the larger social settings in which language contact is found. The social consequences of contact can vary. In some cases, they may appear to be benign or advantageous; in other cases, as Thomason (2001: 5) points out:

> it is not surprising that the social effects of language contact are sometimes painful or even lethal. In many such cases the language of a minority culture is used by a dominant culture as a marker of cultural differentness: not only does it provide a means of identifying the people to be discriminated against, but it also offers a target for discrimination.

Language contact, as all who work in this area are quick to point out, involves social issues, and as Bowern (2010b: 341) writes, "any linguistic claim about language contact reduces to a claim about social behavior of speakers".

Research in contact linguistics can be particularly sensitive to certain aspects of ethical research. For one thing, issues of power, equality, and independence often loom large, as discussed. Not only are there often power differences between the researcher and participants, but language contact situations may also involve internal power differences (as noted in the quote from Thomason above), which can be challenging from an ethical perspective.

Different types of ethical questions might arise, depending on the particulars of the contact situation. For instance, who is an appropriate researcher in a language contact situation? Gal (1979), studying language use in a German-Hungarian community in Austria, first interviewed people in Hungarian, as she was a native speaker of Hungarian. She later interviewed many of the same people together with a monolingual German student, making use of German mandatory. She reflects on the reduced range of language variation that she found in the interviews, and the need to find ways of recording naturally occurring conversation in a wide variety of settings. While Gal's work extended over a long period of time, one might ask how her experience as a speaker of Hungarian (and not of German) might have affected the research and the validity of the conclusions. How would people choose to accommodate to her? When the interviewer was a German-speaking student, would they accommodate her differently? If the researcher were from the community, would things have been different? While it might seem *a priori* prefer-

able to have the researcher come from the community, how might the social roles which they have in that community affect the research?

Another challenge comes in the very definition of 'language'. Matras (2009: 4) stresses that speakers have

> a complex repertoire of linguistic structures at their disposal. This repertoire is not organized in the form of 'languages' or 'language systems'; the latter is a meta-linguistic construct and a label which speakers learn to apply to their patterns of linguistic behaviour as part of a process of linguistic socialisation. Rather, elements of the repertoire […] gradually become associated, through […] linguistic socialisation, with a range of social activities, including factors such as sets of interlocuters, topics, and institutional settings. Mature multilingual speakers face a constant challenge to maintain control over their complex repertoire of forms and structures and to select those forms that are context-appropriate. Context-appropriate selection does not necessarily conform to a separation of 'languages': In some contexts, certain types of cross-linguistic 'mixing' and 'inserting' may be socially acceptable and may constitute effective goal-oriented communication.

Eckert (2013: 12) offers a similar observation, albeit in a somewhat different way: "for most people in the world, language is inextricable from personal and community identity". Language is far more than structure and lexicon: language is used to convey knowledge, social status, power, and identity.

Given this complex sense of language, whereby language systems are viewed as social constructs, an understanding of how the linguistic repertoire is employed is important. If a person's 'languages' are distinguished largely in a 'meta' way, how should situations be identified in which contact effects can be studied? What does it mean to show contact effects if, in some ways, there is only a single language system, with speakers making choices (conscious or not) about how to deploy it? What is the range of factors that must be considered to develop a deep understanding of contact at a point in time? Depending on the research questions at issue, different considerations will be in focus, yet it is important to keep in mind that research questions have implications for research methods, i.e. where ethics must be considered. It is all too easy to think in terms of methods involving surreptitious recordings, circumspection about goals, and attempts to control in artificial ways how people make choices about what part of their linguistic repertoire from which to draw.

7. Summary

Current approaches to ethics in linguistics foreground issues of shared rights, obligations, and responsibilities, with attention to personal and community obligations being essential, as Trechter (2013: 33) concludes. Trechter (2013: 34–35) also highlights that there may be conflict in obligations to: the public, professional standards, students and colleagues, the needs and desires of language communities, and protection of individual research participants. In addition, research methods beyond informed consent and who is considered to be a 'researcher' are questions of ethical concern. In considering methods and researchers, and what kinds of answers different choices might lead to, it is useful to keep in mind Irvine and Gal's (2000: 36) previously cited statement: "[t]here

is no 'view from nowhere,' no gaze that is not positioned". As Schilling (2013: 265) observes, today's general consensus is that "our human interactions with our human subjects are always more important than our academic gain". Precisely what this means differs depending on circumstances, and this is one of the major challenges of doing ethical fieldwork in contact linguistics. Thinking in local ways, as opposed to in absolutes, is a pathway to doing research that is recognized as 'ethical' in terms of relationships, knowledge, and methods, while also understanding that what is considered 'ethical research' will likely shift over time, just as language does.

8. References

Besnier, Niko
2013 Vignette 3a: Responsibility to research participants in representation. In Christine Mallinson, Becky Childs & Gerard Van Herk (eds.), *Data collection in sociolinguistics*, 46–49. New York and London: Routledge.

Bischoff, Shannon & Carmen Jany (eds.)
2018 *Insights from practices in community-based research: From theory to practice around the globe*. Berlin: de Gruyter Mouton.

Bowern, Claire
2010a Fieldwork and the IRB: A snapshot. *Language* 86(4). 897–905.

Bowern, Claire
2010b Fieldwork in contact situations. In Raymond Hickey (ed.), *The handbook of language contact*, 340–357. Malden, MA: Wiley-Blackwell.

Bowern, Claire
2015 *Linguistic fieldwork: A practical guide*, 2nd edn. Basingstoke: Palgrave Macmillan.

Cameron, Deborah, Elizabeth Frazer, Penelope Harvey, M. B. H. Rampton & Kay Richardson
1992 *Researching language: Issues of power and method*. New York: Routledge.

Chelliah, Shobhana & Willem de Reuse
2011 *Handbook of descriptive linguistic fieldwork*. Dordrecht: Springer.

Czaykowska-Higgins, Ewa
2009 Research models, community engagement, and linguistic fieldwork: Reflections on working within Canadian Indigenous communities. *Language Documentation & Conservation* 3(1). 15–50.

Dobrin, Lise & Josh Berson
2011 Speakers and language documentation. In Peter K. Austin & Julia Sallabank (eds.), *The Cambridge handbook of endangered languages*, 187–211. Cambridge: Cambridge University Press.

Dobrin, Lise & Saul Schwartz
2016 Collaboration or participant observation? Rethinking models of 'linguistic social work'. *Language Documentation & Conservation* 10. 253–277.

Dwyer, Adrienne
2006 Ethics and practicalities of cooperative fieldwork and analysis. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 31–66. Berlin: Mouton de Gruyter.

Eckert, Penelope
2013 Ethics in linguistic research. In Robert Podesva & Devyani Sharma (eds.), *Research methods in linguistics*, 11–26. Cambridge: Cambridge University Press.

Gal, Susan
1979 *Language shift: Social determinants of language change in bilingual Austria*. New York: Academic Press.

Garrett, Paul
2004 Language contact and contact languages. In Alessandro Duranti (ed.), *A companion to linguistic anthropology*, 46–72. Oxford: Blackwell Publishing.
Grenoble, Lenore A. & Simone S. Whitecloud
2014 Conflicting goals, ideologies and beliefs in the field. In Peter K. Austin & Julia Sallabank (eds.), *Beliefs and ideologies in language endangerment, documentation and revitalization*: *Proceedings of the British Academy 199*, 339–356. Oxford: Oxford University Press.
Himmelmann, Nikolaus
2006 Language documentation: What is it and what is it good for? In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 1–30. Berlin: Mouton de Gruyter.
Holton, Gary
2009 Relatively ethical: A comparison of linguistic research paradigms in Alaska and Indonesia. *Language Documentation & Conservation* 3(2). 161–175.
Irvine, Judith T. & Susan Gal
2000 Language ideology and linguistic differentiation. In Paul V. Kroskrity (ed.), *Regimes of language: Ideologies, polities, and identities*, 35–83. Santa Fe: School of American Research Press.
Jaffe, Alexandra
1999 *Ideologies in action: Language politics in Corsica*. Berlin: Mouton de Gruyter.
Krauss, Michael
1992 The world's languages in crisis. *Language* 68(1). 4–10.
Kroskrity, Paul V.
2004 Language ideologies. In Alessandro Duranti (ed.), *A companion to linguistic anthropology*, 496–517. Oxford: Blackwell Publishing.
Kulick, Don
1992 *Language shift and cultural reproduction: Socialization, self, and syncretism in a Papua New Guinean village*. Cambridge: Cambridge University Press.
Kulick, Don
1998 Anger, gender, language shift, and the politics of revelation in a Papua New Guinean village. In Bambi Schieffelin, Kathryn Woolard & Paul Kroskrity (eds.), *Language ideologies: Practice and theory*, 87–102. Oxford: Oxford University Press.
Labov, William
1972 Some principles of linguistic methodology. *Language in Society* 1. 97–120.
Labov, William
1982 Objectivity and commitment in linguistic science: The case of the Black English trial in Ann Arbor. *Language in Society* 11. 165–202.
Leonard, Wesley Y. & Erin Haynes
2010 Making 'collaboration' collaborative: An examination of perspectives that frame field research. *Language Documentation & Conservation* 4. 268–293.
Macaulay, Monica
2012 [2004] Training linguistics students for the realities of fieldwork (reprint). In Nicholas Thieberger (ed.), *The Oxford handbook of linguistic fieldwork*, 457–472. Oxford: Oxford University Press.
Matras, Yaron
2005 Language contact, language endangerment, and the role of the 'salvation linguist'. In Peter K. Austin (ed.), *Language documentation and description*, vol. 3, 225–251. London: SOAS.
Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.

Moore, Robert E., Sari Pietikäinen & Jan Blommaert
2010 Counting the losses: Numbers as the language of language endangerment. *Sociolinguistic Studies* 4(1). 1–26.
National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research
1979 *Belmont Report: Ethical principles and guidelines for the protection of human subjects of research*. Washington: US Department of Health, Education, and Welfare.
Nuremberg Code
1947 https://history.nih.gov/research/downloads/nuremberg.pdf. (accessed on 12 September 2018).
Rice, Keren
2006 Ethical issues in linguistic fieldwork: An overview. *Journal of Academic Ethics* 4. 123–155.
Rice, Keren
2009 Must there be two solitudes? Language activists and linguists working together. In Jon Reyhner & Louise Lockard (eds.), *Stabilizing indigenous languages: Encouragement, guidance, and lessons learned*, 37–59. Flagstaff, Arizona: Northern Arizona University.
Rice, Keren
2010 The linguist's responsibilities to the community of speakers: Community-based research. In Lenore A. Grenoble & N. Louanna Furbee (eds.), *Language documentation: Practice and values*, 25–36. Amsterdam: John Benjamins.
Rice, Keren
2011 Documentary linguistics and community relations. *Language Documentation & Conservation* 5. 187–207.
Rice, Keren
2012 Ethical issues in linguistic fieldwork. In Nicholas Thieberger (ed.), *The Oxford handbook of linguistic fieldwork*, 407–429. Oxford: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199571888.013.0019. (accessed on 12 September 2018).
Rumsey, Alan
2010 Ethics, language, and human sociality. In Michael Lambek (ed.), *Ordinary ethics: Anthropology, language, and action*, 104–111. New York: Fordham University Press.
Sallabank, Julia
2013 *Endangered languages: Identities and policies*. Cambridge: Cambridge University Press.
Schilling, Natalie
2013 *Sociolinguistic fieldwork*. Cambridge: Cambridge University Press.
Secretariat on Responsible Conduct of Research
2014 *Tri-Council policy statement: Ethical conduct for research involving humans*. Ottawa: Government of Canada. http://www.pre.ethics.gc.ca/pdf/eng/tcps2-2014/TCPS_2_FINAL_Web.pdf. (accessed on 12 September 2018).
Shuster, Evelyne
1997 Fifty years later: The significance of the Nuremberg Code. *The New England Journal of Medicine* 337(20). 1436–1440.
Thomason, Sarah
2001 *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
Trechter, Sara
2013 Social ethics for sociolinguistics. In Christine Mallinson, Becky Childs & Gerard Van Herk (eds.), *Data collection in sociolinguistics: Methods and applications*, 33–45. New York: Routledge.
WMA
1964–2013 *WMA declaration of Helsinki: Ethical principles for medical research involving human subjects*. https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/. (accessed on 12 September 2018).

Wolfram, Walt
1993 Ethical considerations in language awareness programs. Paper presented at the American Association of Applied Linguistics Ethics Symposium, Atlanta, Georgia, April 1993. https://files.eric.ed.gov/fulltext/ED361998.pdf. (accessed on 12 September 2018).
Wolfram, Walt
2013 Community commitment and responsibility. In J. K. Chambers & Natalie Schilling (eds.), *The handbook of language variation and change*, 2nd edn., 557–578. Oxford: John Wiley & Sons.

Keren Rice, Toronto (Canada)

V. Interactions with neighboring disciplines

55. Anthropology

1. Anthropology in language contact studies
2. Language contact phenomena in anthropological research
3. The relationship between language and culture: linguistic relativity
4. Identities in language contact research
5. Language ideologies
6. Ethnographic methods
7. Future directions
8. References

1. Anthropology in language contact studies

The study of language contact in anthropology is characterized by a focus on how linguistic forms are part of the social lives of speakers. While detailed linguistic description and theories of language structure are certainly also part of an anthropological approach, a given language is studied as cultural behavior within the societies in which it is spoken (Garrett 2004). Languages in contact are thus not solely structures which may converge or intertwine, but also communicative performances which develop as socially situated practices. The emphasis is not on the code *per se*, but on the interactions in which the code is used and through which it develops. Research questions address not just what the structures are and/or which linguistic models and theories we can use to describe and explain them, but also why these particular structures are found in their unique cultural and societal contexts, developing at particular points in history.

In the next section of this chapter, I briefly outline different types of language contact phenomena and how they have been studied within anthropology, and then follow with a section on the relationship between language and culture. The final two sections address contemporary foci in language contact research in anthropology, social identities, and language ideologies, and I end with sections on methodology and future directions.

2. Language contact phenomena in anthropological research

The goal of language documentation led anthropologists to study minoritized languages (often in postcolonial contexts), which are not the dominant languages in the societies where they are spoken. Language documentation is often an integral part of linguistic revitalization (Grenoble and Whaley 2006; Hinton and Hale 2001). Increasingly, such projects seek to do more than describe languages: they also address issues of language ideology (a topic which will be discussed in more detail in Section 5 below; see for example Meek 2011). Hinton (2010) discusses how researchers (working in both the fields of linguistics and anthropology) can contribute to these endeavors in ethical ways, acknowledging the self-empowerment of indigenous populations and the importance of the societally dominant language in their lives, while nonetheless supporting language acquisition, preservation, and literacy in the indigenous language.

https://doi.org/10.1515/9783110435351-055

Although some of these languages being documented are isolated and spoken by monolinguals, in most cases, these languages are in contact with majority languages (and perhaps other minority languages as well), hence the desire for revitalization. In the latter case, the speakers of these languages are multilinguals. These language contact situations give rise to research questions concerning multilingual repertoires, codes, and patterns of use, and the related issue of factors guiding language maintenance and shift. Anthropological approaches to these topics move beyond documentation to examine a broad range of language contact phenomena as part of societal practices.

In many language contact situations, minority languages are not necessarily maintained, and the study of language maintenance and shift addresses how social factors contribute to patterns of language use (or loss). For example, Gal (1978) presents a study of how marriage patterns are motivated by the social status of members of different linguistic groups, leading to language shift among German-Hungarian bilinguals in a small town in Austria. However, as discussed by Croft (2000), exolinguistic exogamy may lead to various linguistic consequences: in some cases, very little borrowing, in others, heavy lexical (and some morphological) borrowing. Moving beyond the influence of the institution of marriage on language use, Errington (1998) addresses the role of state institutions in language shift in Indonesia, looking at the political context, as well as the social identities of speakers, and how these influence their linguistic performance, which in turn determine over time the vitality of languages in a community.

In some cases, language shift can mean more than just the loss of a language within a particular community, and can actually mark the end of the language as a mother tongue as a whole. Dorian (1989) is a landmark work on language death, providing case studies which examine in their social contexts the linguistic mechanisms of obsolescence. More recent research emphasizes the interactional nature of the linguistic process of language death. For example, Ahlers (2006) addresses how the use of Native American languages by nonnative speakers can be used to frame English speech events as part of the construction of a Native American identity. Henne-Ochoa and Bauman (2015) address the performance of generational identity as part of the societal process of language obsolescence in a Lakota community. In these studies, the linguistic structure of dying languages is linked to the interactional motivations of speakers embedded in societal norms and institutional practices.

In maintenance of bilingualism, diglossia is a key concept. Ferguson (1959) presents a situation in which there is a one-to-one correspondence between language and domain. In Ferguson's work, as well as subsequent research by Fishman (1967), diglossia was presented as the main pattern through which multilingualism can be maintained. However, while this concept, and its focus on the societal norms surrounding language choice, is central to the anthropological study of language contact, it is far from the majority pattern in contact situations. Most societies are not diglossic, and there is a great deal of research on what has been called code-switching, or the juxtaposition of two codes in a single conversation. A seminal work on this topic by Gumperz and Blom (1972) launched investigation into the social meanings of language choice: speakers do not simply speak the language normatively associated with a social situation, but instead use their languages in creative ways. Blom and Gumperz (1972) used the term 'metaphorical code-switching' to describe the use of a language to index different social roles and societal meanings. Heller (1988) discussed language choice in terms of shared community norms and negotiation of the social positioning of individuals, that is, while particular

speakers may use language in unique ways, these linguistic performances are nonetheless linked to societal values around specific languages and the people associated with them. In Section 4 below, I discuss more recent research on code-switching as part of the construction of social identities.

In addition to the maintenance and use of two (or more) codes in situations of multilingualism, contact varieties may also emerge. These phenomena have been studied under the rubrics of convergence, creolization, and mixed languages. A landmark study by Gumperz and Wilson (1971) examined language use by inhabitants of a village in India, Kupwar, where varieties of Mararathi, Kannada, and Urdu were spoken. They describe a situation of convergence, wherein linguistic distinctiveness was maintained, but the languages developed word-for-word transability. While the process shared some features with creolization (discussed in the next paragraph), it did not entirely share either the social scenario or the linguistic outcome typical of creole development. Instead, the languages converged toward each other, showing contact-induced features while also not merging entirely into a single code. This topic of convergence has been further developed by Thomason and Kaufman (1988), who presented a correlation between the intensity of cultural contact and the resultant types of lexical and structural borrowing. While their work has attracted criticism (see for instance Sankoff [2001] and Thomason [2008] for a discussion of these critiques), this work motivated research on linguistic convergence as a societally motivated and interactionally driven process.

Related to the study of convergence is the study of mixed languages, or languages which have been traditionally described as deriving much of their grammar from one language and their lexicon from another; however, this is an oversimplification of the complex patterns of language intertwining (Bakker and Muysken 1994; Matras and Bakker 2003) found in languages such as Michif (a French-Cree mixed language), Ma'a/Mbugu (a language with elements from the Bantu and Cushitic language families), and Media Lengua (Quechua-Spanish). Research within anthropology has examined language mixtures, such as Urban Wolof (Swigart 1992, 1994, 2000) and Town Bemba (Spitulnik 1999), with a focus not on their structures, but rather on the language ideologies and social meanings associated with these codes. Swigart (2000) illustrates how Urban Wolof, spoken in Dakar, Senegal, although lacking the official status of French, has nonetheless acquired in some spheres (i.e. the informal economy and popular media) an 'alternative legitimacy' and power in these interactions. Spitulnik (1999) discusses Town Bemba, which she describes as encompassing myriad multilingual, Bemba-based linguistic practices. The use of this code represents an act of identity, with speakers using mixed varieties to index urbanity within and outside of urban centers.

Approaches to language contact phenomena which focus on identity are also found in the study of pidgin and creole languages. Le Page and Tabouret-Keller (1985) made the notion of 'language as an act of identity' a focus in language contact research with their study of the Creole-speaking Caribbean and West Indian communities of London. The central tenet of this work was that individuals use language to associate themselves with various social groups, and to distance themselves from others. This work moved the analysis of language structure and patterns of use away from a focus on social norms and domains as determinants of how speakers use language, and toward a focus on speaker agency. Anthropology has by now long embraced poststructuralist and social constructivist ways of looking at language and identities (a topic taken up in Section 4),

which has had an important influence on language contact studies (see Chapter 24 for a more extensive discussion of the role of identity in language contact research).

3. The relationship between language and culture: linguistic relativity

Another focus within linguistic anthropology which has contributed to the field of language contact research is the study of the relationship between language and worldview. Studies of linguistic relativity address the issue of whether, and how, language influences thought. According to some views, speakers of different languages perceive the world around them in different ways; most well known is the Sapir-Whorf Hypothesis (see Kay and Kempton [1984] for a discussion of this). Inherent to this perspective is the position that we must study multiple languages in order to understand the peoples of the world. Although the study of language and worldview does not necessitate research on language contact *per se*, it does bring to the fore comparative as well as contact studies. In a well-known study on this, Lucy (1996) compared nominal number marking in Yucatec Maya and American English, arguing that the different linguistic systems in these two languages correlate with different priorities in the perception of nouns (with English speakers paying more attention to number and shape, and Yucatec Maya speakers showing more focus on the material composition of objects). Scholars continue to debate this issue of connections between language, culture, and cognition; for example, Deutscher (2010) argues for attention to the importance of the conditioning of our thought processes through language and culture, while McWhorter (2014) claims that language does not shape worldview (however, see a critical review of the latter in Webster [2015]).

Despite the disputed nature of the connections between language and culture, the assumption that there is indeed a connection is of course central to linguistic anthropology. This is seen, for example, in the study of language ideologies (as discussed in Section 5 below), a strand of linguistic anthropology which has become an important aspect of language contact studies (see Ball [2012] and Leavitt [2011] for further discussion).

4. Identities in language contact research

Within poststructuralist approaches, social identities are seen as the outcome, and not the source, of language use. Identities are thus socially constructed, multiple, and fluid (Bucholtz and Hall 2005). Of particular interest here is the concept of authenticity (or 'enoughness' [Blommaert and Varis 2011]) when positioning a given speaker in a given identity category. Being a speaker of (a) particular language(s) is viewed not as a static fact, but rather as part of a linguistic performance. Both the social categorization of 'being a speaker of X' as well as the borders of the associated membership group are negotiated through discourse; not just the social identities of speakers, but also the identity categories themselves, are thus socially constructed through language use. This perspective is especially relevant to studies on language choice, and has been used to look at how speakers use particular languages to index particular aspects of identity, especially

ethnic and/or national identities, as well as gender and sexuality, social class, and/or age. For instance, Fuller (2007, 2012) investigates the identities of preteens in bilingual programs in Berlin, Germany, and southern Illinois, USA, noting that language choice is part of the construction not just of ethnic and national identities, but also of gender and social class group belonging. Similarly, Mendoza-Denton (2008) presents an ethnography of two Latina gangs and how they construct their distinctive identities through language and embodied practices: it is because of shared ethnic commonalities that the distinctions between the two groups are critical to the construction of the identities of the girls as members of different gangs, which are linked to racial and social class markers.

In some cases, the identities of speakers are contested and must be negotiated. Bailey's (2001) work is particularly illustrative of how language is part of the social construction of racial categories, with the Dominican Americans in his study identified by others as 'Black' or 'Spanish' based on their use of English or Spanish, respectively. Negotiation of identity as a speaker of a particular language takes on a different sort of importance for language learners. In this language contact context, the issue of identities changing over time is particularly salient: as individuals become speakers of another language, they may in that process begin to identify differently, in particular, as belonging to their adopted country and culture, while resisting being continually constructed as the 'Other'. However, such shifts in identification may not be readily recognized and accepted by members of the dominant group, and thus must be continually renegotiated (Warriner 2007).

A further topic in the study of linguistic construction of social identities (and which is central to the study of language contact) is the examination of national identities. A national language is often seen as a symbol of national belonging; this can be seen in analyses of conversation and language choice. For example, Shenk (2007) discusses a negotiation of the identity of 'Mexican' in the discourse of three Mexican-American college students, which includes among other characteristics proficiency in Spanish. This connection between language and nation is taken up again in Section 5 in the discussion of language ideologies and policies.

The link between identity and language also plays a role in macrostudies of language in society, where identification with a minority linguistic group is critical in language maintenance and revitalization. On the one hand, language policies supporting minority languages can be motivated by ethnic group solidarity (García 2012); on the other hand, the daily use of a language is also stronger when it is seen as a key element of identity (Jaffe 2007). However, it is important to note not only that identities are multiple and shifting, but also that patterns of language use entail more than just a choice between speaking either one language or the other. As discussed above, there are various linguistic use patterns (e.g. code-switching, convergence) that may occur. For example, Rampton ([2005] 2014) presents an ethnographic study of youth language in multicultural London neighborhoods, illustrating the practice of 'crossing', or the use of codes that index social categories to which one does not belong. Rampton's data show youths of different backgrounds using bits of Panjabi and Creole in addition to various forms of English, including the stylized Asian English spoken by some of their parents, local dialects, and formal registers. The social meanings of these codes are simultaneously drawn on and developed through interactions among peers.

Glick-Schiller, Basch, and Blanc-Szanton (1992) introduced the term 'transnationalism' to describe the situation in which immigrants have connections, experiences, and identities linked to both their country of origin and their country of settlement. They go back and forth across borders (often both literally and figuratively), and have social networks which connect them to both societies. The identities which are constructed through transnational lives are linked to more than one nation − and often more than one language − and in this way are hybrid, because they are a challenge to essentialist ideas about members of these categories. That is, some categories, such as nationality, ethnicity, or gender, are viewed as internally homogenous and mutually exclusive. As such, individuals who identify as being members of more than one category within the set go against widespread ideologies about the nature of these social categories. Bilingual discourse, or any sort of speech viewed as mixing elements from two linguistic systems symbolically linked to two identity categories, is a common index of hybrid identity (Woolard 1998b); the study of transnationalism thus necessarily addresses the study of language contact.

Theoretically, the concept of hybridity problematizes cultural boundaries constructed through essentialist discourses (Smith and Leavy 2008), and can create what has been called a 'third space' (Bhabha 1994). Although the third space is not solely about language, the concept can be applied to multilingual settings where speakers construct a space in which the mixing of codes is part of in-group practices.

Another term rooted in the experiences of migration and multilingualism is 'superdiversity' (Blommaert and Rampton 2011; Blommaert and Varis 2011; Vertovec 2007, 2010), which is a 'diversification of diversity' due to increased migration and subsequent changes in populations. In addition to the proliferation of cultural and linguistic contact, increased use of mobile communication technology also creates networks that are larger and more diverse. The linguistic consequences of this are many. In some cases, it creates higher and higher linguistic demands on migrants, and also contributes to more fractured and fluid identities of speakers. While there have been criticisms of this term, challenging both the newness of the situation and the innovativeness of the concept (Pavlenko 2018), this line of research is part of a growing depiction of what has been called 'fractured multilingualism' (Graber 2015). In this line of inquiry, we move away from viewing languages as discrete, bounded entities, in favor of viewing linguistic production as the result of speakers drawing on all of the available resources in their linguistic repertoires (Makoni and Pennycook 2005; Rymes 2014). Much of the research resulting from this perspective on multilingualism examines the role of language in geographical and social mobility, as discussed above in terms of superdiversity. Language use indexes social identities, and through new ways of combining and using languages, creates new social categories across space and time.

While identity plays a role in these issues, language ideologies are also interwoven with similar societal aspects of language contact. In the next section, the anthropological underpinnings of research on language ideologies are discussed.

5. Language ideologies

Another key topic in linguistic anthropology which relates to language contact is the study of language ideologies. Language ideologies are ideas about language structure

and use which index the political and economic interests of individuals, and the social groups and nations to which they belong (Errington 2000). Although there is some overlap between the study of language attitudes and language ideologies (see Chapters 20 and 25, respectively), the study of ideologies retains a focus on societal discourses, and is based on poststructuralist ideas about how linguistic behavior shapes social realities. Woolard (1992) discusses four central features of the study of language ideologies. First, such research focuses on belief systems and their implications, centered on how these function in society (rather than in the minds of individuals). Second, while derived from the positions and perspectives of a subset of the society under study, language ideologies are often viewed as universally true; we can thus talk about hegemonic ideologies. Third, ideologies are often discussed as rationalizations, distortions, or false positions, which is closely linked to the fourth feature, i.e. that they are inherently linked to social power and legitimation. The study of language ideologies thus involves critical perspectives on both discourses about language and how such discourses inform language choice, which in turn influences language maintenance and shift in multilingual communities. Particularly in the study of endangered languages, where language documentation has long been part of the agenda in linguistic anthropology, the focus on ideologies has brought a critical perspective to previously descriptive endeavors (Johnson and Ricento 2013).

Of particular focus in language contact studies are monoglossic versus pluralist ideologies (Fuller 2019). Monoglossic ideologies value one language − the socially dominant language, which may or may not be spoken by the majority of the population − above all others within a particular society. This leads to the devaluing both of multilingualism, and of linguistic codes other than the socially dominant one. Pluralist ideologies value multilingualism, which usually means fostering maintenance and encouraging use of minority as well as majority languages. However, this does not necessarily mean that multilingual practices are encouraged or even tolerated, with the result that, even within pluralist ideologies, stigmatization of multilingual discourse may still occur (Blommaert et al. 2012).

Furthermore, standard language ideologies often prevail under both monoglossic and pluralist discourses. Gal (2012) illustrates this through her discussion of language and diversity in the European Union (EU), tracing the history of hegemonic monolingualism and showing how the EU's fostering of multilingualism is not a change in ideology, but merely the replacement of the hegemony of one standard language in one nation with the hegemony of multiple standard languages in a set of united nations. This multilingualism is not all-inclusive, instead creating new hierarchies which still exclude speakers of particular languages. Zentz (2014) further discusses language ideologies in Indonesia, illustrating again how sociolinguistic hierarchies are created under overtly pluralist ideological stances. Language policies urging people to 'love their local language', but to speak the national language, espouse the value of local languages while simultaneously delegitimatizing them for use.

This focus on inequalities also extends to studies of language ideologies in transnationalism. Dick (2011) discusses how languages are not neutral indices of social groups, but rather index stratified sets of people, some of which are positioned as 'belonging' in a particular national context, while others are not. In the US for instance, the English-only ideology positions speakers of other languages as backward, un-American, and irrational for not wanting to be part of the American Dream. Multilingualism is thus framed as inherently problematic. This links back to the discussion above about

transnational identities, which challenge such essentialist understandings of social categories.

Finally, language ideologies about the value and roles of specific languages are critical aspects of language contact settings, and often key factors in the continued viability of said languages. Meek (2007) exemplifies this in her research on a Northern Athapascan community, showing that the heritage language, Kaska, has become ideologically compartmentalized as part of the traditional indigenous knowledge which is the domain of the elders. Since Kaska is linked to the position of authority of the elders, it is a code which is inappropriate for children to use. As a result, despite the positive value of the language, young members of the community are not socialized to use it in their daily lives; these ideologies thus contribute to language shift rather than maintenance.

Yet ideological dominance is never complete (Woolard 1998a): there are always competing ideologies. Proctor (2014) writes about different ideologies surrounding English in India. Both English-dominant and English-excluding ideologies focus on the benefits of English in terms of social mobility, while also creating different views of social hierarchy in Indian society as linked to English proficiency. The English-dominant discourse frames English as a commodity that can be acquired by anyone, while the English-excluding discourse recognizes that there are barriers to the acquisition of English and its cultural capital. The English-dominant discourse thus rejects the effect of privilege, framing language learning as a result of personal effort, as opposed to an outcome of a lifestyle that is not attainable for all. Inherent to both of these perspectives is the ideology of the instrumental value of a language.

Such studies of language ideologies are a prime example of the need for methodological approaches which provide insights into not just descriptions of linguistic forms, but also the societal and interactional factors of community language use. In the next section, I present a brief discussion of the role of ethnography in language contact research.

6. Ethnographic methods

A key influence from the field of anthropology has been the use of ethnography to understand the role of language in a particular culture. The principle of studying the social functions of different linguistic forms, derived from ethnography of communication (Hymes 1964, 1994), is a guiding perspective in much research today incorporating poststructuralist social theories.

Beyond the method of ethnography of communication, ethnographic methods more generally (which often include interviews, recording of speech events, surveys, questionnaires, and/or document analyses) are the mainstay of anthropological research. This reliance on qualitative research to incorporate the broader cultural and societal context into analyses of the intricacies of language use defines the contribution of anthropology to the study of language contact.

The value of ethnography in the study of language contact can be seen in many of the above mentioned works. For instance in Meek (2007), without ethnographic work which looked at the broader cultural practices of both adults and children, the avoidance of speaking Kaska could easily be analyzed as a lack of interest in traditional ways, as opposed to an adherence to heritage culture values. Mendoza-Denton's (2008) ethno-

graphic research situates language as part of a more complex system of social categorization and identity construction. Swigart's (2000) research on Urban Wolof relies on an in-depth understanding of both linguistic practices and their tacit social values; how such values develop in communities of practice can be seen in other ethnographic work, such as Bailey (2001) and Rampton (2014), which is a central contribution of anthropology to the study of language contact.

7. Future directions

Topics in language contact research which will continue to benefit from anthropological perspectives include new media and multimodal analysis (Akkaya 2014; Reyes 2014), as well as the closely related issues of immigration and transnationalism, and of intertextuality and interdiscursivity. Just as people move through time and space, words move from one context to another via various modalities. Computer and new media discourse create ever more possibilities for language contact; anthropology can contribute to the analysis of the social meanings and linguistic forms created. Androutsopoulos (2008) addresses the potential of online ethnographies as part of his research on computer-mediated discourse and other new media data. As in other contexts, participant observation can be indispensable in contributing to emic understandings of cultural behavior. Another arena in which different modalities are part of language contact research is the study of multilingual linguistic landscapes (Blommaert 2013); for a comprehensive discussion of this topic, see Chapter 36 of this volume.

The contributions from anthropologists, and all scholars who draw on anthropological treatments of languages in contact, can guide us in our quest to describe and interpret linguistic forms and uses. More importantly they help us to challenge static ideologies and discriminatory practices in contexts of language contact.

8. References

Ahlers, Jocelyn C.
2006 Framing discourse. *Journal of Linguistic Anthropology* 16(1). 58–75.

Akkaya, Aslihan
2014 Language, discourse, and new media: A linguistic anthropological perspective. *Language and Linguistics Compass* 8(7). 285–300.

Androutsopoulos, Jannis
2008 Potentials and limitations of discourse-centred online ethnography. *Language@internet* 5(8). http://www.languageatinternet.org/articles/2008/1610. (accessed on 13 September 2018)

Bailey, Benjamin
2001 The language of multiple identities among Dominican Americans. *Journal of Linguistic Anthropology* 10(2). 190–223.

Bakker, Peter & Pieter Muysken
1994 Mixed languages and language intertwining. In Jacques Arends, Pieter Muysken & Norval Smith (eds.), *Pidgins and creoles: An introduction*, 41–52. Amsterdam: John Benjamins Publishing Company.

Ball, Christopher
2012 Boasian legacies in linguistic anthropology: A centenary review of 2011. *American Anthropologist* 114(2). 203–216.
Bhabha, Homi K.
1994 *The location of culture*. London: Routledge.
Blom, Jan-Petter and John. J. Gumperz
1972 Social meaning in linguistic structure: Code-switching in Norway. In Jan-Petter Blom, John J. Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics*, 407–434. New York: Holt, Rinehart and Winston.
Blommaert, Jan
2013 *Ethnography, superdiversity and linguistic landscapes: Chronicles of complexity*. Bristol: Multilingual Matters.
Blommaert, Jan & Ben Rampton
2011 Language and superdiversity. *Diversities* 13(2). 1–22.
Blommaert, Jan & Piia Varis
2011 Enough is enough: The heuristics of authenticity in superdiversity. *Tilburg Papers in Culture Studies* 2. 1–13.
Blommaert, Jan, Sirpa Leppänen, Päivi Pahta & Tiina Räisänen (eds.)
2012 *Dangerous multilingualism: Northern perspectives on order, purity and normality*. Basingstoke: Palgrave Macmillan.
Bucholtz, Mary & Kira Hall
2005 Identity and interaction: A sociocultural linguistic approach. *Discourse Studies* 7(4–5). 585–614.
Croft, William
2000 *Explaining language change: An evolutionary approach*. London: Longman.
Deutscher, Guy
2010 *Through the language glass: Why the world looks different in other languages*. New York: Metropolitan books.
Dick, Hillary Parsons
2011 Language and migration to the United States. *Annual Review of Anthropology* 40. 227–240.
Dorian, Nancy C. (ed.)
1989 *Investigating obsolescence: Studies in language contraction and death*. Cambridge, UK: Cambridge University Press.
Errington, Joseph
1998 On the state of a language of state. In Bambi B. Schieffelin, Kathryn A. Woolard & Paul V. Kroskrity (eds.), *Language ideologies: Practice and theory*, 271–284. Oxford University Press.
Errington, Joseph
2000 Ideology. *Journal of Linguistic Anthropology* 9(1–2). 115–117.
Ferguson, Charles A.
1959 Diglossia. *Word* 15(2). 325–340.
Fishman, Joshua A.
1967 Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues* 23(2). 29–38.
Fuller, Janet M.
2007 Language choice as a means of shaping identity. *Journal of Linguistic Anthropology* 17(1). 105–129.
Fuller, Janet M.
2012 *Bilingual pre-teens: Competing ideologies and multiple identities in the US and Germany*. New York: Routledge.

Fuller, Janet M.
2019 Language ideologies. In Lourdes Ortega & Annick De Houwer (eds.), *The Cambridge handbook of bilingualism*, 119–134. Cambridge: Cambridge University Press.
Gal, Susan
1978 Peasant men can't get wives: Language change and sex roles in a bilingual community. *Language in Society* 7(1). 1–16.
Gal, Susan
2012 Sociolinguistic regimes and the management of 'diversity'. In Alexandre Duchêne & Monica Heller (eds.), *Language in late capitalism: Pride and profit*, 22–42. London: Routledge.
García, Ofelia
2012 Ethnic identity and language policy. In Bernard Spolsky (ed.), *The Cambridge handbook of language policy*, 79–99. Cambridge: Cambridge University Press.
Garrett, Paul B.
2004 Language contact and contact languages. In Alessandro Duranti (ed.), *A companion to linguistic anthropology*, 46–72. Oxford: Blackwell Publishing.
Glick-Schiller, Naomi, Linda Basch & Cristina Blanc-Szanton
1992 Transnationalism: A new analytic framework for understanding migration. *Annals of the New York Academy of Sciences* 645(1). 1–24.
Graber, Kathryn E.
2015 On the disassembly line: Linguistic anthropology in 2014. *American Anthropologist* 117(2). 350–363.
Grenoble, Lenore A. & Lindsay J. Whaley
2006 *Saving languages: An introduction to language revitalization*. Cambridge: Cambridge University Press.
Gumperz, John J. & Robert Wilson
1971 Convergence and creolization. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 151–167. Cambridge: Cambridge University Press.
Heller, Monica (ed.)
1988 *Codeswitching: Anthropological and sociolinguistic perspectives*. Berlin: Walter de Gruyter.
Henne-Ochoa, Richard & Richard Bauman
2015 Who is responsible for saving the language? Performing generation in the face of language shift. *Journal of Linguistic Anthropology* 25(2). 128–149.
Hinton, Leanne
2010 Language revitalization in North America and the new direction of linguistics. *Transforming Anthropology* 18(1). 35–41.
Hinton, Leanne & Kenneth Hale (eds.)
2001 *The green book of language revitalization in practice*. Boston, MA: Brill.
Hymes, Dell
1964 Introduction: Toward ethnographies of communication. *American Anthropologist* 66(6[2]). 1–34.
Hymes, Dell
1994 Toward ethnographies of communication. In Janet Maybin (ed.), *Language and literacy in social practice: A reader*, 11–22. Clevedon: Multilingual Matters.
Jaffe, Alexandra
2007 Minority language movements. In Monica Heller (ed.), *Bilingualism: A social approach*, 50–70. London: Palgrave Macmillan.
Johnson, David C. & Thomas Ricento
2013 Conceptual and theoretical perspectives in language planning and policy: Situating the ethnography of language policy. *International Journal for the Sociology of Language* 219. 7–21.

Kay, Paul & Willet Kempton
1984 What is the Sapir-Whorf hypothesis? *American Anthropologist* 86(1). 65–79.
Leavitt, John
2011 *Linguistic relativities: Language diversity and modern thought*. Cambridge: Cambridge University Press.
Le Page, Robert B. & Andrée Tabouret-Keller
1985 *Acts of identity: A creole based study of language and ethnicity*. Cambridge: Cambridge University Press.
Lucy, John A.
1996 *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
Makoni, Sinfree & Alastair Pennycook
2005 Disinventing and reconstituting languages. In Sinfree Makoni & Alastair Pennycook (eds.), *Disinventing and reconstituting languages*, 1–41. Clevedon, UK: Multilingual Matters.
Matras, Yaron & Peter Bakker (eds.)
2003 *The mixed language debate: Theoretical and empirical advances*. Berlin: Mouton de Gruyter.
McWhorter, John H.
2014 *The language hoax: Why the world looks the same in any language*. New York: Oxford University Press.
Meek, Barbra A.
2007 Respecting the language of elders: Ideological shift and linguistic discontinuity in a Northern Athapascan community. *Journal of Linguistic Anthropology* 17(1). 23–43.
Meek, Barbra A.
2011 *We are our language: An ethnography of language revitalization in a Northern Athabascan community*. Tucson: University of Arizona Press.
Mendoza-Denton, Norma
2008 *Home girls: Language and cultural practice among Latina youth gangs*. Malden, MA: Blackwell.
Pavlenko, Aneta
2018 Superdiversity and why it isn't. In Stephan Breidbach, Lutz Küster & Barbara Schmenk (eds.), *Sloganizations in language education discourse*, 142–168. Bristol, UK: Multilingual Matters.
Proctor, Lavanya M.
2014 English and globalization in India: The fractal nature of discourse. *Journal of Linguistic Anthropology* 24(3). 294–314.
Rampton, Ben
2014 [2005] *Crossings: Language and ethnicity among adolescents*. London: Routledge.
Reyes, Angela
2014 Linguistic anthropology in 2013: Super-new-big. *American Anthropologist* 116(2). 366–378.
Rymes, Betsy
2014 *Communicating beyond language: Everyday encounters with diversity*. New York: Routledge.
Sankoff, Gillian
2001 Linguistic outcomes of language contact. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 638–668. Oxford: Blackwell.
Shenk, Petra Scott
2007 'I'm Mexican, remember?' Constructing ethnic identities via authenticating discourse. *Journal of Sociolinguistics* 11(2). 194–220.

Smith, Keri E. Iyall & Patricia Leavy (eds.)
2008 *Hybrid identities: Theoretical and empirical examinations*. Boston, MA: Brill.
Spitulnik, Debra
1999 The language of the city: Town Bemba as urban hybridity. *Journal of Linguistic Anthropology* 8(1). 30–59.
Swigart, Leigh
1992 Two codes or one? The insiders' view and the description of codeswitching in Dakar. *Journal of Multilingual & Multicultural Development* 13(1–2). 83–102.
Swigart, Leigh
1994 Cultural creolisation and language use in post-colonial Africa: The case of Senegal. *Africa* 64(2). 175–189.
Swigart, Leigh
2000 The limits of legitimacy: Language ideology and shift in contemporary Senegal. *Journal of Linguistic Anthropology* 10(1). 90–130.
Thomason, Sarah G.
2008 Social and linguistic factors as predictors of contact-induced change. *Journal of Language Contact* 2(1). 42–56.
Thomason, Sarah G. & Terrence Kaufman
1988 *Language contact, creolization and genetic linguistics*. Berkeley, CA: University of California Press.
Vertovec, Steven
2007 Super-diversity and its implications. *Ethnic and Racial Studies* 30(6). 1024–1054.
Vertovec, Steven
2010 Towards post-multiculturalism? Changing communities, conditions and contexts of diversity. *International Social Science Journal* 61(199). 83–95.
Warriner, Doris S.
2007 Language learning and the politics of belonging: Sudanese women refugees becoming and being 'American'. *Anthropology & Education Quarterly* 38(4). 343–359.
Webster, Anthony K.
2015 Why the world doesn't sound the same in any language and why that might matter: A review of *The language hoax: Why the world looks the same in any language* by John H. McWhorter. *Journal of Linguistic Anthropology* 25(1). 87–93.
Woolard, Kathryn A.
1992 Language ideology: Issues and approaches. *Pragmatics* 2(3). 235–249.
Woolard, Kathryn A.
1998a Introduction: Language ideology as a field of inquiry. In Bambi B. Schieffelin, Kathryn A. Woolard & Paul V. Kroskrity (eds.), *Language ideologies: Practice and theory*, 3–47. Oxford, UK: Oxford University Press.
Woolard, Kathryn A.
1998b Simultaneity and bivalency as strategies in bilingualism. *Journal of Linguistic Anthropology* 8(1). 3–29.
Zentz, Lauren
2014 "Love" the local, "use" the national, "study" the foreign: Shifting Javanese language ecologies in (post-)modernity, postcoloniality, and globalization. *Journal of Linguistic Anthropology* 24(3). 339–359.

Janet M. Fuller, Groningen (The Netherlands)

56. Colonial studies

1. Introduction
2. Colonial studies and linguistics
3. Factors determining colonial language contact and its documentation
4. Language contact in colonial settings
5. Conclusion
6. References

1. Introduction

Language was and is an important part of colonial structures and of the exertion of colonial power. In multiple ways, colonial power structures had an influence on language ecologies, language practices, repertoires, and communication patterns, as well as on the conception of the interrelationship between languages. Colonialism is one of the most prominent language-external causes of language contact. It was also the conceptual framework for the description of many of the languages that came into contact through colonialism. Errington (2001: 19–20) therefore admonishes that "'[c]olonialism' is a rubric for hugely different exploitative purposes, institutional configurations, and modes of subordination; so the work of linguistic description done under the aegis of various colonial regimes needs to be considered with an eye to conditions that enabled it and social interests inscribed in it".

Terminologically, colonialism is understood here to mean the unilaterally claimed political dominance of one societal entity over another, noncontingent one (Warnke, Stolz, and Schmidt-Brücken 2016). This includes modern-age (i.e. postmedieval) colonial activities, such as the colonization of Ibero-America and the actions of the imperial nations of the 19th century claiming possession of overseas territories and carrying out a political, military, and cultural rule of dominance over a population that was considered subordinate in power and cultural development.

This chapter takes a closer look at the interaction(s) between colonialism and language, and explores the interrelations between colonial studies and language contact studies. The second section introduces the topics of language and colonialism, and of colonial linguistics. It highlights why colonial and postcolonial aspects are relevant for language contact studies. The third section identifies factors that contribute to the specific conditions and outcomes of colonial language contact. Colonial power constellations interact with linguistic transfer and influence; considering both contact linguistic and colonial specific factors leads to a better understanding of these processes. The fourth section is concerned with the interaction of these factors. Based on examples from different colonial language contact settings, the impact(s) of relevant factors as well as typical linguistic outcomes are detailed. The final section concludes by summarizing sociolinguistic patterns that result from language contact under colonial conditions.

2. Colonial studies and linguistics

Colonial and *postcolonial studies* originated as subdisciplines of history and literature studies, but have since developed into independent and multifaceted disciplines. They

https://doi.org/10.1515/9783110435351-056

are concerned with research-based investigation, analysis, and critical assessment of colonialism and postcolonialism. As complex phenomena, colonialism and postcolonialism require multidisciplinary approaches; one such approach is that of *colonial and postcolonial linguistics*.

Colonial and postcolonial linguistics refers to the interrelationship between colonialism and language, and between colonialism and the historical development of linguistics as a discipline. Calvet (1974) and Errington (2001) were seminal in shaping engagement with language in colonial contexts as a subject of research. Calvet (1974) sharply demonstrates how colonial power structures contribute to the rescaling and death of languages; Errington (2001) shows how the development of (comparative) linguistics, and in certain aspects, even modern linguistic approaches, are rooted deeply in colonial thinking patterns.

One research area within colonial and postcolonial linguistics is the study of language contact that occurred under colonial conditions. Colonialism changed language ecologies and instigated contact between originally noncontiguous languages. Colonizer languages were transferred to colonized areas, resulting in mutual linguistic influence as well as in the development of new languages and varieties (Mühlhäusler 1996).

Regarding the language system, languages are affected in similar ways by both colonial and noncolonial language contact. Nevertheless, the sociolinguistic conditions and ideological framing of the two settings are clearly different, justifying a specific colonial studies approach to colonially induced language contact. The colonial setting entails a structure of power, oppression, and imposition that is not found in other contact settings in the same way (Calvet 1974). It determines what languages come into contact, how languages are perceived and ranked, what is done to increase or diminish contact between them, and in what ways the outcomes of such contact are judged, supported, stigmatized, or suppressed.

3. Factors determining colonial language contact and its documentation

The wide range of colonial encounters resulted in many different contact scenarios. Typical settings of colonial interaction were (colonial) administration, church and Christianizing, education (governmental, private, and mission schools; boarding and day schools), (para)military settings (including the movement of paramilitary groups to various parts of the colonial empire, see Morlang 2008), trade, and colonial work settings (indented labor in plantations, local work in mission station workshops, colonially administered work such as building roads, and so on). There were considerable differences in the local application of colonial language policies, for example, with respect to schooling and language instruction, influencing the intensity of contact as well as the effects of such contact. Furthermore, selecting a local language as the mission or church language of a wider area led to its more widespread adoption. This in turn resulted in notable numbers of language learners. Among the observed outcomes were koineïzation effects, which are linguistic levelling processes between largely mutually intelligible varieties resulting in the loss of strongly distinctive features and the emergence of a

compromise variety. For an example of this kind of colonially induced language change, see Ross (1996) on the case of Kâte in Papua New Guinea.

Linguistic factors (in a narrower sense) affect the different subsystems of language. The outcome is shaped by specific language contact constellations leading to (mainly lexical) borrowing and transfer versus (mainly structural) imposition effects (van Coetsem 2000), or language maintenance versus language shift. Lexical borrowing tends to be more typical of language maintenance settings, while structural changes are found more often in language shift and language learner scenarios (Thomason and Kaufman 1988). In colonial language contact, these linguistic factors interact with the colonial power asymmetry.

The specifically colonial setup impacted not only the conditions of colonial language contact and its outcomes, but also its documentation. Three different areas deserve special attention in this respect.

3.1. Language attitudes and ideologies

As an ideologically charged endeavor, colonialism shaped the way social structures, cultures, and languages were perceived and interpreted. The power asymmetry between the colonial homeland and the colonized areas was replicated in language attitudes and in judgements about colonizer versus colonized languages. This had an effect on the intensity of contact, and on the evaluation of language contact effects (as strong or mild). Power asymmetry also played a role in whether or not contact varieties were classified as languages or were considered appropriate research objects (e.g. pidgins and creoles).

3.2. The data base: data collection

The documentation of colonized languages was usually carried out by nonnative speakers, including by missionaries (who could be fluent speakers), but also by European-origin ethnographic researchers and natural scientists who had no previous command of these languages. Often (but not always), they relied on the help of (few) local speakers or on translators. They were frequently not able to judge the quality of the data they collected (see Kayser [1917] for a very critical assessment of Hambruch [1914] along these lines).

3.3. The data base: transferring spoken languages into writing

Colonial-time language contact is, to a large extent, a diachronic topic, and the underlying processes can often only be deduced from the preserved documents of that time. The colonial-time documentation of languages frequently entailed colonial agents (e.g. missionaries) designing a graphemic system and putting languages into writing. Which varieties were considered languages, and which forms of communication were actually transferred into a written code, was influenced among other factors by ideological con-

siderations of language purism and linguistic unification, where a range of spoken varieties was reduced to one single written version (Paris 2012; Ross 1996; Spolsky 2009).

In addition, the societal position of the (relatively few) colonizers allowed for lasting influence on the colonized languages, even though they were nonnative speakers of the local languages: their learner varieties were, in some cases, assigned a special status, e.g. as the Christian ritual variety or as the codified written standard. Examples are a learner variety of Kâte in New Guinea (Paris 2012), or the historical variety of Guaraní misionero in Paraguay (Melia 2003).

So when interpreting colonial-time object-language and metalinguistic data, it is important to consider: (a) which language was documented, and (b) who did the documentation. These aspects heavily influenced the representation of languages and their usage patterns (Errington 2001, 2008; Lüpke and Storch 2013; Wolff 2016).

4. Language contact in colonial settings

The mutual influence between colonizer and colonized languages impacted the linguistic systems of the languages involved and led to language change. This is frequently observed in the lexicon, but depending on the type of contact, structure can be affected, too.

It should be noted that the documented version of the language may have diverged from the everyday spoken version even at the time of data collection. Apparent changes may have resulted from undetected misunderstandings (e.g. with respect to phonetics, semantics, or the function of specific morphemes), or from the informants' consciously withholding information (e.g. with respect to spiritual topics). Under these circumstances, it can be difficult to establish a 'baseline' against which to measure potential changes and contact phenomena. Any research on contact-induced change in these languages requires a careful comparison of all available language materials, combined with meta- and extralinguistic information, and a critical investigation of the actual historical contact and documentation setting.

4.1. Contact effects on colonized languages

4.1.1. Lexical effects

Contact-induced change in colonized languages is most easily detectable in the lexicon, where borrowed items can be identified based on sound and meaning correspondences. Lexical documentation, however, was often marked by descriptive selectivity that led to an artificially 'purified' version of a variety which arbitrarily ignored (older) borrowings that were integral parts of its lexicon. For example, Hambruch (1914) only listed nonborrowed words of Nauruan, leaving out loanwords from South Sea Pidgin English that are included in other compilations from the same era (e.g. Delaporte 1907; Kayser [1936] 1993). In a similar vein, Christaller, a German governmental teacher in Cameroon who put Duala into writing, explicitly notes that he chose a variety that seemed to him to be the 'purest' one in terms of lexical borrowing (Adick and Mehnert 2001: 246).

While such ideological considerations veiled lexical interference in some cases, there are other instances where high rates of foreign vocabulary were imposed on colonized languages. This practice was most pervasive with respect to Christian vocabulary (underscoring Tadmor's [2009: 64] claim on the common nature of borrowing in the world's languages in the semantic field of religion). The efforts of missionaries to fill assumed lexical gaps with foreign words and to reframe spiritual terminology are instances of lexical and semantic language planning (cf. Zimmermann [2005] for Spanish terminology in Nahuatl texts in Mexico, and Stolberg [2017] for German terminology in Bel texts in New Guinea). Other areas were affected as well, however, such as work settings and administration (Mühlhäusler 2012). Loanwords are thus partly side effects of colonialism, linked to the introduction of foreign concepts, tools, and practices ('unplanned' language planning, see Baldauf 1994), and partly the result of language policies favoring colonizer languages in work and mission environments. It is not always clear whether all of the imposed loanwords were actually used by native speakers of the language. In some cases, however, later documentation confirms the preservation of such loanwords (Engelberg, Möhrs, and Stolberg 2017; Groves 1938; Mühlhäusler 1977; van der Sijs 2015).

Given the low number of native speakers of colonizer languages in colonized areas, the overall amount of borrowing is quite high, and it becomes obvious that the power constellation of colonialism must have considerably reinforced the impact of borrowing. On the other hand, case studies such as that of Engelberg (2006) show that the contact outcome was also mediated by additional factors, such as the previous presence of one colonizer language (e.g. English) that could block borrowing from another colonizer language (e.g. German).

4.1.2. Morphological effects

While direct morphological transfer is rare, some morphological effects of colonial language contact have been reported. Among them is the introduction of a grammatical gender system in Chamorro, an Austronesian language under heavy colonial influence from Spanish (Stolz 2012a). Another instance is found in Otomí (Mexico), where verbal conjugation patterns were replaced by Spanish prepositions due to colonial language contact; similar changes are attested for other Amerindian languages as well (Suaréz 1983; Zimmermann 1992).

In some cases, however, apparent morphological change seems to be a matter of (biased) interpretation, as Stolz (2012b) and Hackmack (2012, 2015) show. Examples are the colonial-time descriptions of word separation in Chamorro or of case distinctions in Swahili. In these instances, the (European-based) linguistic expectations of the non-native language users who documented these languages seem to have influenced their judgments of the languages, with the result that morphological patterns were rendered inadequately.

4.1.3. Graphemic effects

Many colonized languages were oral languages. Missionaries, in particular, devised graphemic systems and put the languages they worked with into writing in order to provide

converts with access to the Bible. Often, a body of Christian literature was produced as well (Bible translations, hymn and prayer books, Bible stories for children, etc.). When introducing a writing system, the missionaries tended to use phoneme-grapheme correspondences from their native languages. One example is the lengthening 'h' used in Pohnpeian to mark the length of a preceding vowel, transferred from German orthography (Pohnpeian spelling was originally developed by German missionaries; see Dalby 1998; Rehg 2004). Distinctions between sounds and their phonetic/phonemic status were also not always recorded properly. While the phonology itself was usually not affected, the orthographic systems of formerly colonized languages thus still bear the traces of this graphemic transfer. This resulted in inconsistent or competing spelling systems in several cases, with negative consequences for effective schooling (e.g. in Nauru and Pohnpei; see Lotherington 1998; Rehg 2004).

4.2. Contact effects on colonizer languages

4.2.1. Lexical effects

The extent of linguistic effects on the colonizer languages varied for different colonial settings. Initially, the effects were mainly lexical, and they remained so in short-term colonial settings. In the case of German, for example, the documented lexical influence is fairly limited, due in part to the short duration of the German colonial period (circa 30 years) and to processes of substitution and attrition after contact with German had ended (after WWI).

In long-term colonial settings, structural effects are attested as well. Ibero-American Spanish, with a more than 500 year history of colonization in America, shows considerable lexical borrowing (mainly of content words) from American languages. Some of this influence has carried over to European Spanish. In general, however, lexical transference as well as structural contact phenomena remained restricted to the varieties that developed in the actual contact settings, and they had little noticeable effect on the standard language of the colonial metropole (i.e. the homeland of the colonial power). This can be attributed to language ideological attitudes towards colonial varieties of colonizer languages: the so-called indigenized varieties of colonizer languages, such as Ibero-American Spanish, were not viewed as being on a par with their European counterparts (Zimmermann 2015).

4.2.2. Structural effects and new varieties

Where colonizer languages were taught and spread, structural interference occurred in the speech of language learners and nonnative speakers. On rare occasions, learner texts in the colonizer languages are preserved, such as students' texts from missionary and government schools (e.g. Dwucet 1908). They reflect learner errors to an expected (if limited) degree; these features often remained restricted to the learner register, however, and did not necessarily lead to a newly evolving contact language.

In many cases, however, structural interference from native (colonized) language(s) onto nonnative (colonizer) language(s) led to the development of new varieties, such as pidgin or creole languages, with the colonizer language as the main lexifier and one or more colonized languages contributing structural features; this largely occurred in multicontact settings, such as plantations (see Thomason and Kaufman [1988]). They either stabilized to varying extents, resulting in a wide range of European-lexified pidgins and creoles, or they were spoken during, and for some time beyond, the colonial period, dropping out of use after the actual contact had ended (see Deumert [2009] on Namibian Black German; Maitz [2016] and Volker [1991] on Rabaul Creole German; and Mühlhäusler [1977, 2012] on Pidgin German in Oceania).

These new varieties were often negatively evaluated (see Engelberg 2014; Huber and Velupillai 2018; see Kutzner [2012] for similar attitudes regarding North American languages). Such attitudes about colonial languages were frequently grounded in an ideologically based justification of colonial imperialism as well as in economic interests (see Migge and Léglise [2007] for British and French colonies).

It has been debated whether pidgins and creoles that draw the main part of their lexicon from a colonial lexifier language should be considered varieties of this lexifier language. Mufwene (2009, 2015) argues that the same contact and change processes underlie the development of, for example, North American English and English-based creoles in Africa or Asia. He maintains that typological distinctions between the various outcomes of colonial language contact are based on socioideological grounds, not on linguistic ones. DeGraff (2005), in addition, discusses some socially problematic implications of considering creoles to be linguistically exceptional. From this perspective, colonizer languages have undergone structural change to substantial degrees.

5. Conclusion

Language contact under colonial circumstances was influenced by historical conditions, political decisions, and the duration of colonial dominance. Colonizer languages as well as colonized languages were affected to different degrees. New varieties, languages, and patterns of communication evolved, while others were extinguished or are endangered as a result of colonialism.

It is hardly possible to exhaustively name the regions and countries impacted by colonialism, because it has been, and continues to be, so vast a phenomenon. Modern-time colonial domination exerted by (mainly) European powers began in the late 15th century and continued into the 20th century (and in some cases, into the 21st century). The study of (post)colonial linguistics extends to all these cases, including Africa, the Americas, Asia, Australia, Oceania, and some parts of Europe, with differing colonial constellations across time.

In terms of language ecology, language spread, and language death, a number of characteristic outcomes of colonial domination and colonially induced language contact are attested. Zimmermann (2015: 3–4) offers an overview of Ibero-America; other colonial language contact settings have yielded comparable results. They include the following scenarios.

Indigenous and local languages:
- Indigenous (and other local) languages were often exterminated or marginalized, that is, reduced to geographically limited areas, limited domains of use, or limited use by specific social groups.
- Indigenous languages – as far as they still exist(ed) – were and are influenced structurally and sociostructurally by the (former) colonizer languages as a result of planned and unplanned effects of colonial (language) politics.
- Some indigenous languages, e.g. Quechua (South America) and Kâte (Papua New Guinea), were spread to other areas through colonialism and missionaries' activities, resulting in new contact and learner scenarios.

Colonizer languages:
- Formerly colonial languages were indigenized, that is, the colonial language became (one of) the dominant language(s) of the indigenous population and showed structural effects of shift-induced language contact. These indigenized colonial languages are often still not recognized as standard varieties (e.g. Ibero-American Spanish), or they have come under pressure from subsequent imperial languages (e.g. Spanish *vis-à-vis* English in the southwestern USA; see Riegelhaupt, Carrasco, and Brandt 2003; Zimmermann 2015).
- Some colonizer languages came into contact with indigenous languages displaced by forced migration, relocation, and/or slavery (e.g. Spanish and Portuguese were in contact with sub-Saharan languages in South America).
- In several cases, subsequent colonizer languages competed with and replaced each other (e.g. German replaced Spanish in the Mariana Islands; English replaced French in Louisiana and parts of Canada; French replaced Chinese in Vietnam; Arabic replaced French which had replaced Arabic in the Maghreb; see Alalou 2006).

Contact varieties:
- Mixed languages, pidgins, and creole languages evolved as a result of colonial language contact.

From a postcolonial perspective, Spolsky (2004) summarizes the outcomes of colonial language contact and language policies into three categories of postcolonial language communities:

a) original/indigenous languages were erased; the former colonizer's language has become the dominant (or only) official language (e.g. Australia, Canada, New Zealand, USA);
b) original/indigenous languages still play some role, side by side with former colonizer languages (e.g. in a number of African and South American countries, such as Cameroon, Niger, Nigeria, Bolivia, and Paraguay); and
c) original/indigenous languages are the primary language(s) of the postcolonial state (e.g. Samoa).

Colonialism gave rise to language contact in many ways, forced and deliberate. The movement and dislocation of people(s), often across long distances, led to new contact settings. Speakers' interactions were framed by the power asymmetry and political conditions of colonialism. Language documentation, both in colonizing and in colonized languages, was usually carried out by agents of the colonizing power. As a result, there is still a colonially induced bias in the available language data. To adequately analyze these

data not only requires an informed language contact perspective, but also a thorough understanding of the specific colonial conditions that applied during the production of the data.

Combining language contact studies and colonial studies can result in a better understanding of both disciplines' research topics. The colonial and postcolonial perspective on language contact highlights the importance of specific environmental conditions: commonly observed effects may be compromised due to sociopolitical factors that are specific to colonial settings. Examples of such factors are colonial language policies, missionary language planning, and the effects of colonial and postcolonial language ideologies. Contact-induced change may happen when speakers of different languages interact, but it may also be suspended, even under seemingly favorable conditions (see Thomason 2000). The language contact settings that can be found in the context of colonial studies provide rich evidence of both outcomes.

6. References

Adick, Christel & Wolfgang Mehnert
2001 *Deutsche Missions- und Kolonialpädagogik in Dokumenten: Eine kommentierte Quellensammlung aus den Afrikabeständen deutschsprachiger Archive 1884−1914*. Frankfurt/Main: IKO-Verlag für Interkulturelle Kommunikation.

Alalou, Ali
2006 Language and ideology in the Maghreb: Francophonie and other languages. *The French Review* 80(2). 408−421.

Baldauf, Richard B., Jr
1994 "Unplanned" language policy and planning. *Annual Review of Applied Linguistics* 14. 82−89.

Calvet, Louis-Jean
1974 *Linguistique et colonialisme. Petit traité de glottophagie*. Paris: Payot.

Dalby, Andrew
1998 *Dictionary of languages*. London: A & C Black Publishers.

DeGraff, Michel
2005 Linguists' most dangerous myth: The fallacy of creole exceptionalism. *Language in Society* 34(4). 533−591.

Delaporte, Philip A.
1907 *Kleines Taschenwörterbuch: Deutsch-Nauru*. Nauru: Missions-Druckerei.

Deumert, Ana
2009 Namibian Kiche Duits: The making (and decline) of a neo-African language. *Journal of German Linguistics* 21(4). 349−417.

Dwucet, F.
1908 Als Lehrer und Bienenzüchter auf Saipan. In Josef Wiese (ed.), *Neu-Deutschland: Unsere Kolonien in Wort und Bild*, 363−368. Berlin: Meidinger's Jugendschriften Verlag. http://marshall.csu.edu.au/CNMI/CNMIBIB/0545.pdf. (accessed on 18 July 2018).

Engelberg, Stefan
2006 The influence of German on the lexicon of Palauan and Kosraean. In Keith Allen (ed.), *Selected Papers from the 2005 Conference of the Australian Linguistic Society*. http://www.als.asn.au/proceedings/als2005/engelberg-german.pdf. (accessed on 18 July 2018).

Engelberg, Stefan
2014 Die deutsche Sprache und der Kolonialismus: Zur Rolle von Sprachideologemen und Spracheinstellungen in sprachenpolitischen Argumentationen. In Heidrun Kämper, Peter

Haslinger & Thomas Raithel (eds.), *Demokratiegeschichte als Zäsurgeschichte: Diskurse der frühen Weimarer Republik*, 307–332. Berlin & Boston: de Gruyter.

Engelberg, Stefan, Christine Möhrs & Doris Stolberg
2017 Wortschatz deutschen Ursprungs in Tok Pisin. In Peter Meyer & Stefan Engelberg (eds.), *Lehnwortportal Deutsch*. Mannheim: Institut für Deutsche Sprache. http://lwp.ids-mannheim.de/dict/tokpisin. (accessed on 18 July 2018).

Errington, James J.
2001 Colonial linguistics. *Annual Review of Anthropology* 30. 19–39.

Errington, James J.
2008 *Linguistics in a colonial world: A story of language, meaning, and power*. Malden & Oxford: Blackwell.

Groves, William Charles
1938 *Grammar of the language of Nauru*. Pacific Manuscripts Bureau (PMB 1164, Reel 5). Australia National University: Menzies Library.

Hackmack, Susanne
2012 Die Subjektpräfixe des Swahili in kolonialzeitlichen Sprachbeschreibungen. In Stefan Engelberg & Doris Stolberg (eds.), *Kolonialzeitlicher Sprachkontakt: Sprachliche Begegnungen und Auseinandersetzungen*, 49–67. Berlin: Akademie Verlag.

Hackmack, Susanne
2015 Case in selected grammars of Swahili. In Klaus Zimmermann & Birte Kellermeier-Rehbein (eds.), *Colonialism and missionary linguistics*, 51–74. Berlin, München & Boston: de Gruyter.

Hambruch, Paul
1914 *Die Sprache von Nauru (Verwaltungsbezirk der Marshallinseln)*. Hamburg: Friedrichsen & Co.

Huber, Magnus & Viveka Velupillai
2018 Die Database of Early Pidgin and Creole Texts: Sprachplanung, Sprachideologien und Sprachattituden gegenüber dem Pidginenglisch in Deutsch-Neuguinea. In Birte Kellermeier-Rehbein, Matthias Schulz & Doris Stolberg (eds.), *Sprache und (Post)Kolonialismus: Linguistische und interdisziplinäre Aspekte*, 127–160. Berlin/Boston: de Gruyter.

Kayser, P. Alois
1917 Die Eingebornen von Nauru (Südsee): Eine kritische Studie. *Anthropos* 12/13. 313–337.

Kayser, P. Alois
1993 [1936] *Nauru Grammar*. Karl H. Rensch (ed.). Yarralumla: Embassy of the Federal Republic of Germany.

Kutzner, Sandy
2012 "Zivilisierte" und "unzivilisierte" Sprachen: Historische Sprachbewertung und das wirklich Fremde in Sprachen. In Stefan Engelberg & Doris Stolberg (eds.), *Kolonialzeitlicher Sprachkontakt: Sprachliche Begegnungen und Auseinandersetzungen*, 215–229. Berlin: Akademie Verlag.

Lotherington, Heather
1998 Trends and tensions in post-colonial language education in the South Pacific. *International Journal of Bilingual Education and Bilingualism* 1(1). 65–75.

Lüpke, Friederike & Anne Storch
2013 *Repertoires and choices in African languages*. Boston & Berlin: de Gruyter.

Maitz, Péter
2016 Unserdeutsch (Rabaul Creole German): Eine vergessene koloniale Varietät des Deutschen im melanesischen Pazifik. In Alexandra N. Lenz (ed.), *German abroad: Perspektiven der Variationslinguistik, Sprachkontakt- und Mehrsprachigkeitsforschung*, 211–240. Göttingen: V & R unipress.

Melia, Bartomeu
2003 *La lengua guaraní en el Paraguay colonial*. Asunción: CEPAG.

Migge, Bettina & Isabelle Léglise
2007 Language and colonialism: Applied linguistics in the context of creole communities. In Marlis Hellinger & Anne Pauwels (eds.), *Handbook of language and communication: Diversity and change*, 297–338. Berlin/New York: Mouton de Gruyter.
Morlang, Thomas
2008 *Askari und Fitafita: "Farbige" Söldner in den deutschen Kolonien*. Berlin: Christoph Links Verlag.
Mufwene, Salikoko
2009 The indigenization of English in North America. In Thomas Hoffmann & Lucia Siebers (eds.), *World Englishes: Problems, properties, prospects: Selected papers from the 13th IAWE conference*, 353–368. Amsterdam: Benjamins.
Mufwene, Salikoko
2015 Colonization, indigenization, and the differential evolution of English: Some ecological perspectives. *World Englishes* 34(1). 6–21.
Mühlhäusler, Peter
1977 Bemerkungen zum "Pidgin Deutsch" von Neuguinea. In Carol Molony, Helmut Zobl & Wilfried Stölting (eds.), *German in contact with other languages*, 58–70. Kronberg: Scriptor Verlag.
Mühlhäusler, Peter
1996 *Linguistic ecology: Language change and linguistic imperialism in the Pacific Region*. London: Routledge.
Mühlhäusler, Peter
2012 Sprachliche Kontakte in den Missionen auf Deutsch-Neuguinea und die Entstehung eines Pidgin-Deutsch. In Stefan Engelberg & Doris Stolberg (eds.), *Kolonialzeitlicher Sprachkontakt: Sprachliche Begegnungen und Auseinandersetzungen*, 71–100. Berlin: Akademie Verlag.
Paris, Hannah
2012 Sociolinguistic effects of church languages in Morobe Province, Papua New Guinea. *International Journal of the Sociology of Language* 214. 39–66.
Rehg, Kenneth L.
2004 Linguists, literacy, and the law of unintended consequences. *Oceanic Linguistics* 43(2). 498–518.
Riegelhaupt, Florencia, Roberto Luis Carrasco & Elizabeth Brandt
2003 Spanish: A language of indigenous peoples of the Americas. In Jon Reyhner, Octaviana V. Trujillo, Roberto Luis Carrasco & Louise Lockhard (eds.), *Nurturing native languages*, 129–140. Flagstaff, AZ: Northern Arizona University.
Ross, Malcolm
1996 Mission and church languages in Papua New Guinea. In Stephen A. Wurm, Peter Mühlhäusler & Darrell T. Tryon (eds.), *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas*, vol. 2.1., 595–617. Berlin & New York: Mouton de Gruyter.
Spolsky, Bernard
2004 *Language policy*. Cambridge: Cambridge University Press.
Spolsky, Bernard
2009 *Language management*. Cambridge & New York: Cambridge University Press.
Stolberg, Doris
2017 Historical sociolinguistics in colonial New Guinea: The Rhenish mission society in the Astrolabe Bay. *Journal of Historical Sociolinguistics* 3(1). 55–92.
Stolz, Thomas
2012a Survival in a niche: On gender-copy in Chamorro (and sundry languages). In Martine Vanhove, Thomas Stolz, Aina Urdze & Hitomi Otsuka (eds.), *Morphologies in contact*, 93–135. Berlin: Akademie Verlag.

Stolz, Thomas
2012b Über die Wortmacherei, oder: Die Verschiebung der Wortgrenzen in der kolonialzeitlichen Sprachforschung (am Beispiel des Chamorro). In Stefan Engelberg & Doris Stolberg (eds.), *Kolonialzeitlicher Sprachkontakt: Sprachliche Begegnungen und Auseinandersetzungen*, 17–48. Berlin: Akademie Verlag.
Suaréz, Jorge A.
1983 *The Mesoamerican Indian languages: Cambridge language surveys.* Cambridge: Cambridge University Press.
Tadmor, Uri
2009 Loanwords in the world's languages: Findings and results. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 55–75. Berlin: de Gruyter Mouton.
Thomason, Sarah Grey
2000 On the unpredictability of contact effects. *Estudios de Sociolinguistica* 1(1). 173–182.
Thomason, Sarah Grey & Terrence Kaufman
1988 *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
van Coetsem, Frans
2000 *A general and unified theory of the transmission process in language contact*. Heidelberg: Universitätsverlag C. Winter.
van der Sijs, Nicoline
2015 *Uitleenwoordenbank*. http://www.meertens.knaw.nl/uitleenwoordenbank/. (accessed on 18 July 2018).
Volker, Craig
1991 The birth and decline of Rabaul Creole German. *Language and Linguistics in Melanesia* 22. 143–156.
Warnke, Ingo H., Thomas Stolz & Daniel Schmidt-Brücken
2016 Perspektiven der Postcolonial Language Studies. In Thomas Stolz, Ingo H. Warnke & Daniel Schmidt-Brücken (eds.), *Sprache und Kolonialismus: Eine interdisziplinäre Einführung zu Sprache und Kommunikation in kolonialen Kontexten*, 1–25. Berlin & Boston: de Gruyter.
Wolff, H. Ekkehard
2016 *Language and development in Africa: Perceptions, ideologies, and challenges*. New York: Cambridge University Press.
Zimmermann, Klaus
1992 *Sprachkontakt, ethnische Identität und Identitätsbeschädigung. Aspekte der Assimilation der Otomí-Indianer an die hispanophone mexikanische Kultur.* Frankfurt am Main: Vervuert.
Zimmermann, Klaus
2005 Traducción, préstamos y teoría del lenguaje: La práctica transcultural de los lingüistas misioneros en el México del siglo XVI. In Otto Zwartjes & Cristina Altman (eds.), *Missionary linguistics II / Lingüística misionera II: Orthography and phonology: Selected papers from the second International Conference on Missionary Linguistics, São Paulo, 10–13 March 2004*, 107–136. Amsterdam/Philadelphia: Benjamins.
Zimmermann, Klaus
2015 Kolonialismus und Sprachgeschichte in Iberoamerika. In Daniel Schmidt-Brücken, Susanne Schuster, Thomas Stolz, Ingo H. Warnke & Marina Wienberg (eds.), *Kolonialinguistik: Sprache in kolonialen Kontexten*, 1–20. De Gruyter.

Doris Stolberg, Mannheim (Germany)

57. Economics and language contact

1. Introduction
2. Linking up economics with language contact issues
3. Language and economic value: Key findings
4. Beyond the market: A broader view of the value of languages
5. Assessment: Using economics to manage language contact
6. References

1. Introduction

This chapter presents the field of specialization known as 'language economics' or 'economics of language'. A chapter on language economics fits quite naturally into a handbook on language contact, since most research on language economics actually investigates questions that only arise because languages are in contact with each other.

Language economics is defined less by a set of topics than by the fact that it approaches language issues from an economic angle, which this chapter proposes to explain. However, because the object to which this approach is applied remains linguistic, language economics needs to operate in a very interdisciplinary spirit. In particular, it must import several key concepts from the language disciplines. For this reason, this chapter will also highlight along the way several areas where research in contact linguistics can be used to strengthen theoretical and empirical work in language economics.

In Section 2, I present the basics of economic reasoning, emphasizing the continuity between language economics approaches on the one hand, and the study of situations of language contact at the micro-, meso-, and macrolevels on the other hand.

In Section 3, I use the example of language at work, which is a central one in language economics, to illustrate how standard economic concepts, such as 'value' and 'production', can be understood in relation to linguistic variables, primarily language skills. A strong suit of the economic approach is that is does not content itself with the general claim that language skills matter in a person's access to employment and income. It proposes testable causal explanations as well as quantitative estimates of the monetary value of language skills on the labor market.

In Section 4, I turn to an important, but generally under-researched, facet of language economics, namely, the challenge of identifying and measuring certain components of the economic value of language (whether in terms of skills, use, or image) that do not have a directly observable market value. This section also considers conceptual difficulties linked to the passage of time and to the choice of the counterfactual used for assessing value, showing that contact linguistics is in a position to provide valuable insights in the treatment of the associated challenges.

Section 5 briefly summarizes the most important ways in which language economics contributes to our understanding of situations of language contact, and then addresses the contribution of the specialty to matters of linguistic justice, which are of crucial importance for the selection, design, and evaluation of language policies.

https://doi.org/10.1515/9783110435351-057

2. Linking up economics with language contact issues

One relatively easy way to approach language economics is to define it as variants of 'work on language questions by economists' (e.g. Breton 1999; Ginsburgh and Weber 2016; Vaillancourt 1985), but more analytical definitions have also emerged (e.g. Grin 2003, 2016; Zhang and Grenier 2013). They define the specialty in terms of its use of an economic perspective on language questions. This results, in turn, in a focus on three groups of questions:

1. how do linguistic variables influence economic ones;
2. conversely, how do economic variables influence linguistic ones; and
3. how can choices regarding language issues be usefully informed by economic analysis, particularly in the selection and design of language policies?

These three groups of questions all proceed from the core paradigm of economics, whose starting point is the concept of scarcity (Grin 2003). It is because resources are scarce that we need to make choices. In the absence of scarcity, abundance would reign and there would be no need to make choices, let alone study how such choices are made. However, scarcity (whether of money, time, energy, food, the support of fellow human beings, or anything else that can be considered a resource) is a banal reality of human experience. This explains why economics as a discipline has been defined as "the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses" (Robbins 1935: 16). In the same way, language economics generally seeks to approach and explain choices affecting (or affected by) language in terms of how actors use scarce resources, no matter the nature of these resources.

In economics, actors (usually called 'agents') are assumed to make choices under a set of constraints. Some of the choices they make bear upon financial or material questions − for example, when they act as workers, as consumers, or as investors. Other choices bear upon nonmaterial or nonfinancial questions, and involve 'nonmarket' values. The literature also speaks in such cases of 'symbolic' values, for example, in the realms of interpersonal relationships, political activism, or leisure. Actors will generally seek their self-interest (rather than working *against* it), but this basic assumption of economics in no way precludes altruism, since other persons' wellbeing (a spouse, a child, a friend) can easily be built into the economic analysis of actors' choices.

In sum, the economic approach does not focus solely on financial goals and resources. It is considerably more flexible and subtle than some poorly informed commentators often appear to assume. Importantly, looking at language issues through the prism of economics in no way implies a 'neoliberal' perspective, which is quite a different thing. The related accusation of 'commodification' usually rests on untenable logical confusions.

The development of language economics began in the 1960s and 1970s with the study of language-based earnings differentials. This *per se* implies a strong orientation towards questions of contact, since such differentials, by definition, emerge on labor markets where more than one language is present. Earnings differentials can be traced back to various linguistic attributes, like agents' first language, their skills in additional languages, or both. Since then, the specialty has branched out in a many directions and now covers a wide range of topics. The total number of scientific publications in the field of language economics may be said to total around 500 (Gazzola, Grin, and Wickström

2016); several recently published collections of papers in edited volumes bear witness to the vitality of research in the field (Carrère 2016; Gazzola and Wickström 2016; Ginsburgh and Weber 2016).

In addition to its epistemological anchoring in economics, the specialty presents two important, arguably defining features:

- The first feature of language economics is that it retains a somewhat marginal status within economics. This point has been discussed at length elsewhere (in particular, Grin 2003, 2016), but can probably be explained by the field's necessarily interdisciplinary orientation. In some cases, economic discourse about language issues can venture sound propositions without worrying overmuch about questions like 'the nature of language'. This applies, for example, to the evaluation of the labor market premium for skills at a certain level in a certain language. For other topics in language economics, however, economists need to develop a subtler understanding of language. This is often the case in research that treats, say, language use as a dependent variable influenced by economic processes. In such cases, several features of the social, political, and cultural embedding of the processes that determine the use of various languages in contact must be taken into account. For this purpose, economics must refer to concepts developed in other social sciences (economics being, after all, one of them) or in the humanities, and the mainstream of the economics profession is not necessarily at ease with this necessity.
- The second feature of language economics is its strong emphasis on language policy questions, whether in terms of *ex-ante* selection and design, or *ex-post* evaluation. Ultimately, economics is a theory of choice, and its usefulness resides in providing a systemic framework for assessing the pros and cons of competing scenarios. Language policy options are *per se* such scenarios. When assessing them, language economics does not replace, let alone displace, the contributions of other disciplines; on the contrary, it combines with knowledge from other disciplines to produce comparative assessments of language policy alternatives.

Language economics first emerged as a response to actual socioeconomic questions. In Canada, it was used to identify the source of earnings (i.e. labor income) differentials between native speakers of English and native speakers of French in the province of Québec, where over 80 % of the population declares French as a first language. Multivariate analyses showed that earnings differentials were linked to individuals' first language (L1), that is, such differentials remained statistically significant even when other determinants of income (such as education, work experience, or skills in a second language) were included in the earnings equations. Whereas a person's education, experience, and elements of human capital are generally seen as socially acceptable sources of differences in earnings, this is not the case for a person's L1. When such differences are revealed by multivariate analysis, they are generally interpreted as proof of discrimination on ethnolinguistic grounds. These early findings therefore established the relevance of corrective public policies to redress inequalities occurring in this language contact situation.

Similar tools have been used to assess the role of language in explaining the socioeconomic disadvantage of immigrants with low levels of competence in the dominant language of their respective host countries (Chiswick and Miller 2007; Esser 2006). These findings support the (unsurprising) notion that investment in the host country

language is crucial to immigrants' socioeconomic advancement. The same approach has been used later, primarily in European contexts, to estimate the value of investing in *foreign* language skills (that is, skills in languages *other than* the locally dominant or official language[s]); we return to these developments below.

Estimation of language-based earnings differentials serves to explain an economic outcome. In this explanation, language appears as one possible explanatory factor alongside many other ones. The theoretical explanation for these findings is that certain language skills contribute to value creation and are rewarded for this reason. For example, the ability to communicate with (solvable) clients in two or three languages rather than just one is an asset that facilitates the sale of goods and services, and thus generates monetary value added. Even if this type of relation between linguistic variables (e.g. skills) and economic variables (e.g. value added) is likely to hold generally, a more fine-grained understanding of several aspects of this relation may be gained by explicitly including knowledge about the conditions of contact between languages. For example, the (often policy-induced) relegitimization of certain minority languages in various domains engenders demand for skills in these languages, which is reflected in positive rates of return for such skills. The processes respectively studied by economics and applied linguistics should therefore be considered in conjunction with each other.

Since the early 1990s, another line of research in language economics has been developing, focusing on the reverse causal direction: it is intended to account for certain linguistic outcomes, treating economic variables as explanatory variables. For example, an economic process of change in the patterns of international trade can affect the relative importance of various languages in the operations of multinational corporations, altering demand for skills on the labor market. This in turn affects the incentive to learn various languages, with an effect further down the line on a typical object of study for contact linguistics, namely, patterns of foreign language learning, resulting in changed distributions of foreign language skills.

A closely related line of research uses economics as an analytical tool to unravel the workings of this complex interplay itself, without necessarily giving standard economic factors a particularly prominent role in the explanation. Such research is, of course, central to the economic approach to language policy, and developments in this direction find direct applications in evaluations of the effectiveness, cost effectiveness, and fairness of competing language policy scenarios, and more generally, in the selection and design of language policies (Gazzola 2014; Gazzola and Grin 2017; Gazzola and Wickström 2016; Grin and Vaillancourt 2015).

The growing attention devoted to language policy selection and design requires closer cooperation between economists and specialists in the language disciplines. This is why since roughly the mid-2000s language economics can be said to have been moving into a new phase (which one might consider the third in its development), which is marked by increasing and more structured interaction with other disciplines, and by more explicit inclusion of sociolinguistic information in economic modeling and/or econometric work.

3. Language and economic value: key findings

Let us return to the question of 'value', since it is a foundational one in economics more generally, and also in language economics specifically. The 'value' of a minority lan-

guage, the benefits and costs of alternative corporate language policies, and the optimal level of social investment in foreign language instruction are all questions that language economists are often invited to address, yet applying the concept of value to language requires some conceptual precautions.

As a first step, let us distinguish 'market' and 'nonmarket' value. Market values are those that can be 'read off' economic data, such as observed market prices: they capture the material or financial value of something (from tomatoes to cars, and using somewhat more elaborate techniques than simple price observations, linguistic or computer skills). However, nonmarket values (to which we return in the following section) are no less relevant from an economic standpoint. This first distinction corresponds to the two rows in Table 57.1 below.

At the same time, it is important to consider *at what level* or *from whose perspective* value is being examined. We may be concerned with value *for the typical individual* (in which case, we would speak of 'private' value) or with value *for society as a whole* (opening up the question of 'social' value). Obviously, the value of a particular skill will vary across individuals. Data availability determines how far it is possible to refine econometric estimations in order to assess value for different groups of individuals facing different conditions.

Dealing with the distinction between private and social value raises major conceptual and methodological problems stemming from the very nature of language. Language is not only used for communication; it also serves to position oneself in social interaction. It is, at the same time, a prime vehicle for the creation, expression, and reproduction of culture. Therefore, it is bound up with individual and group identity, and some functions of language can only be enacted and experienced collectively. Language can also be said to be a type of network, and the value to a particular speaker of knowing a given language changes depending (among other things) on others' decisions to learn or not to learn it. As a result, social value is not just the sum of the private values accruing to individuals. This feature makes estimation very difficult, and at the time of writing, the problem of *aggregating* private values to estimate social value has not, in the case of language, been adequately resolved in scientific research.

Tackling these difficult questions of identification and measurement requires an understanding of linguistic processes with which economics is not properly equipped to deal, but where contact linguistics is of course strategically placed. Bearing these caveats in mind, the distinction between various manifestations of value remains conceptually sound and operational, and can be represented using a two-way table (Table 57.1).

The estimation of market values is relatively straightforward, if adequately detailed, representative data are available. Few countries have such data, and fewer still collect them on a regular basis. Estimates based on representative survey data in Switzerland, which is characterized by sharply demarcated language regions where German, French,

Tab. 57.1: Four types of value

	Private	Social
Market	A	C
Nonmarket	B	D

Tab. 57.2: Social rates of return on second or foreign languages (%) Switzerland, men, 1994–95

L2 → Language region ↓	French	German	English
French-speaking region	–	6.5	4.7
German-speaking region	10.0	–	12.6
Italian-speaking region	21.5	11.7	*ns*

Control variables: education (years), experience (years), experience squared.
Source: Grin (1999: 124). The private earnings differentials used as a basis for the calculations of social rates of return are significant at the 5 % level for all reported coefficients.

or Italian dominate (the Romansh-speaking areas finding themselves in a far less comfortable situation, not discussed here), indicate that competence in German as a foreign language is highly rewarded on the labor market in French-speaking Switzerland, and vice versa. Competence in English is also profitable in all the language regions of the country. The rates of return on such language skills for the male subsample, controlling for education and work experience, fall in the 10 % to 18 % range: these figures, which exemplify private market value (Cell A in Table 57.1), can be interpreted as the additional labor income, in percentage terms, accruing to people who have acquired (approximately) a B2 level in the foreign language concerned, net of the effect of other standard determinants of labor income (Grin, Sfreddo, and Vaillancourt 2010).

We can move from private to social value (Cell C in Table 57.1) by combining the econometric results on wage premiums, such as those reported in the preceding paragraph, with data on spending on second/foreign language teaching and learning. Such data are difficult to come by, and estimates of social rates of return are rarely calculated because education systems typically do not have figures on spending *by school subject*, and estimates have to be derived from other school statistics. The resulting social rates of return (Table 57.2) confirm that, given the amount the Swiss spend on language teaching and learning, foreign or second language skills constitute a highly profitable investment proposition for society as a whole.

In a multilingual context, contact necessarily occurs, although the actual extent of such contact may vary. Therefore, investing in skills that facilitate this contact, allowing various forms of value to emerge as a consequence, generally represents an economically wise allocation of resources for society as well as for individuals. Consider also that, in the case of Switzerland, English as a foreign or second language does not necessarily guarantee the highest rates of social return. Of course, the lowest figure in the table, at 4.7 %, still represents a very attractive return in real terms, but the results also show that, even if we focus on narrowly defined monetary benefits, English is not the only foreign language investment worth considering in a non-English speaking country.

Extensions to the foregoing results, as well as estimates of other manifestations of economic value, are available in the literature. For example, the macroeconomic counterpart of the above figures (more specifically, the share of Switzerland's GDP that can be traced back to residents' second/foreign language skills) is in the region of 9 % to 10 %. This high rate reflects, among other things, the fact that in the Swiss economy, the use of several languages is woven into productive activity itself (as distinct from, say, the

presence of these languages in casual conversations during breaks). This occurs, for example, for domestic trade across internal language boundaries. Such figures constitute the macroeconomic counterpart of the microeconomic processes that result in the profitability of individual plurilingualism. Similar calculations for Québec, with a main focus on competences in English and French, yield estimates in the 3 % to 4 % range.

Although language contact can take place with varying degrees of conflict, economic research shows that multilingualism generally remains associated with positive economic value. The quantitative empirical evidence available is concentrated on a limited number of cases, since, as pointed out before, the full range of data required to perform econometric estimations is seldom collected. However, the available results tell us a consistent story (see Gazzola, Grin, and Wickström [2016] for an extensive review). Apart from giving access to higher earnings and resulting in higher GDP, foreign language skills generally improve access to employment, and all else being equal, the likelihood of keeping one's job when wages go up and firms try to reduce staff costs is higher among bilinguals than monolinguals. The fact that language skills usually have a positive effect on earnings must be understood in a *ceteris paribus* context, implying a multivariate approach. While it is frequently observed that members of nondominant language communities often have below-average gross earnings, the net effect of language is usually nonnegative, indicating that economic disadvantage is due to other factors, such as lower levels of education in the community concerned. Exceptions exist, revealed by the existence of statistically significant negative coefficients for L1; such cases are indicative of ethnolinguistic discrimination (e.g. Chiswick, Patrinos, and Hurst 2007).

These results mostly apply to major languages, where skills in those languages, through one channel or another, contribute to individual and collective prosperity; this holds for several of them (see however, estimates of positive rates of return for a minority language like Irish in Borooah, Dineen, and Lynch [2009]). It may be useful to point out that such findings are not restricted to English only. Thus, the oft-heard pronouncement that 'English is the language of business', which reflects the correct, if informal, observation that English is overproportionately used as the default language between trading partners who do not share the same mother tongue is, incorrectly, taken to imply that it is the *only* foreign language that leads to prosperity. While its usefulness and relevance are undisputed, it is not the only linguistic asset worth investing in.

Another example can be found in the development economics literature, where linguistic diversity in third-world countries has been blamed for sluggish development. However, the use of more elaborate econometric tools suggests that this correlation is a spurious one, and that in fact, linguistic diversity (or 'fragmentation', as some authors call it) in itself generally has no negative impact on GDP per capita (Arcand and Grin 2013). What some authors in development economics have ignored is that in language contact, there are many ways to weave societal multilingualism into everyday life, including economic activity. Thanks to the combination of various strategies such as second or third language acquisition, the use of receptive skills, and translation, linguistic diversity need not be disruptive.

4. Beyond the market: a broader view of the value of languages

Our discussion so far has focused on market value, if only because little is known about the order of magnitude of nonmarket (or 'symbolic') value. The latter, however, is no

less relevant from an economic standpoint because it is a legitimate component of the comparative assessment of alternative situations. In other words, nonmarket effects, whether positive or negative, ought to be taken into account. However, such a complete assessment, to my knowledge, has never been performed because of the difficulty of rigorously defining, identifying, and measuring nonmarket benefits and costs, and hence of determining the corresponding net value obtained by subtracting costs from nonmarket benefits. Economists lack most of the tools that are *a priori* required for identifying and measuring these nonmarket effects. Some evaluation techniques could be imported from environmental economics, since the nonmarket value of languages bears important similarities with environmental value. For example, an important component of nonmarket environmental value is the pleasure one may derive from admiring unspoiled landscapes or wading in clean streams. In the realm of foreign language skills, nonmarket value may proceed from the enjoyment of easier access to foreign cultures and to the people with whom direct communication becomes possible. However, the parallel between linguistic diversity and environmental quality does not necessarily generate ready-made solutions. More focused theoretical and empirical work is required, and contact linguistics is well placed to provide greatly needed concepts for this purpose.

This case for taking a broad view of the value of language(s) and linking economic questions about value with findings from sociolinguistics (in a broad sense) is further supported if we also think about other facets of the problem. Let us consider in particular the issues of 'time' and 'scope', since taking these dimensions into account alters our understanding of the very meaning of value.

'Time' refers to the time horizon considered. The standard earnings equations used to estimate the rates of return presented above cannot tell us whether private earnings differentials and social rates of return estimated at time t will still hold n periods from now, at time $t + n$. On this matter, we only have circumstantial evidence. The latter suggests that as skills in a given language become more widespread, they also become more banal. Consequently, they command an ever-decreasing premium in those sectors of the economy where they were previously highly profitable (with other sectors in which they never were particularly useful remaining unaffected). In other words, even if a particular language is highly useful in many professions, it is likely that with the passage of time and the banalization of skills in that language, it is skills in *other* languages that will give people, and the economies in which they work, a true edge. In order to unravel the dynamics that result in individual choices to use a language or not, to learn it or not, and to pass it on to the next generation or not (presented here as binary, but which of course can be approached as a continuum), and to ascertain the aggregate impact of such choices on macrolevel patterns of language spread and decline, we can bank on theoretical models (e.g. Templin et al. 2016), but these necessarily rest on stylized sociolinguistic processes. Here again, interdisciplinary cooperation can notably improve the appropriateness of the hypotheses adopted in this stylization.

Let us now turn to 'scope', which concerns a very fundamental point in the debate about languages and diversity, albeit one that is often the victim of obfuscation. Understanding the issue of 'scope' requires us to make a basic distinction: on the one hand, we can ask ourselves what foreign language investment is worthwhile, *given that we live in a multilingual world*. This can be termed the 'contingent' question, because it makes sense under a given contingency: specifically, the linguistic diversity that characterizes the world, and with which we are confronted with increasing frequency as a result of

globalization. So far, this paper has been framed in this 'contingent' case. Against this backdrop, teaching and learning languages, English included, is economically valuable. On the other hand, we may ask ourselves what is the value of having linguistic diversity at all − the implicit counterfactual being, in this case, linguistic uniformity. This is what may be called the 'absolute' question, because it raises the very different issue of the value of diversity in the absolute (Grin, Sfreddo, and Vaillancourt 2010). Proving that diversity is valuable from a contingent perspective does not suffice to prove that it is valuable from an absolute perspective: mastering several languages is profitable in the world as it is now, but on balance, might not humankind be better off if it did not need to deal with diversity? Might not prosperity be greater with one single worldwide language?

At this point, we reach the current frontiers of language economics. The examination of the value of diversity *in the absolute* is only incipient. Yet here again, available circumstantial evidence suggests that diversity is a solution far superior to uniformity. Let us leave aside the obvious observation that there is enjoyment to be derived from choice, and choice proceeds from the availability of alternatives − of diversity, as it were. This returns us to the question of nonmarket values, and the treatment of this question is best handled using the methods of environmental economics. It is highly plausible (much more, at any rate, than the *opposite* proposition) that through nonmarket effects, diversity generates value, and hence wellbeing. However, even if we confine ourselves to a narrower facet of economic value, namely, the 'market' value appearing in the top row of Table 57.1, another interesting question arises, namely, that of the possible correlation between multilingualism, creativity, and innovation. Psychometric research into the effect of (individual) plurilingualism or (group) multilingualism on the creativity and capacity for innovation of persons and working teams is only just beginning. Available evidence (Fürst and Grin 2017) suggests that linguistic diversity is superior to uniformity because it broadens the range of mental tools available for problem solving and control, and for coming up with novel ideas − a crucial condition of prosperity in postindustrial, knowledge-based economies.

5. Assessment: using economics to manage language contact

The emphasis placed on the value of language skills as a core issue in language economics should be seen as a federating question connecting several areas of application that have direct relevance for language policy selection, design, and evaluation. The economic approach contributes to the theoretical and empirical knowledge base required to address numerous challenges familiar to scholars in contact linguistics. These challenges are encountered in the various situations where society needs to arbitrate between the competing (or at any rate, not fully converging) interests and goals of different groups defined, at least in part, by their linguistic attributes. Examples include, but are not limited to: the selection and design of measures for minority language protection and promotion; the setting of language learning goals in school curricula; the efficient and fair operations of multilingual organizations; the allocation of roles to local languages and languages of wider communication (*lingue franche*) in higher education; the regulation of linguistic landscapes; and/or the linguistic requirements placed upon allophone immigrants in integration policies.

Readers may note that no mention has been made here of 'commodification', a notion which refers to the treatment of various goods, services, other nonmaterial entities, and even people that were hitherto *not* viewed as such, as potentially tradable 'commodities'. This construct, which has roots in Marxian economic theory, makes sense in that specific theoretical context, but is used sparingly, if at all, in contemporary economics, and is only referred to in some strands (usually those least economically oriented) of the related field of political economy (and it must not be confused, of course, with the process of commoditization in finance). The notion of commodification has limited relevance to language economics for several reasons, the main one being that it usually lacks a clear analytical content. Its policy implications are problematic, too, as shown by the following example: suppose that a minority language community, in order to improve the long-term prospects of the language, democratically chooses to encourage its use in business and commerce, or to market language-specific goods and services. This strategy may perhaps be criticized as involving 'commodification', but such a claim, which typically comes from commentators ensconced in the comfort of a nonthreatened dominant language, raises the uncomfortable question of these commentators' legitimacy in passing this type of judgment on what may be a justifiable, and possibly smart strategic move.

One further point, however, needs to be addressed in order to round out this overview of the specialty and assess its usefulness as a tool that societies can use to handle situations of language contact. Our closing considerations address matters of linguistic justice, which are an essential part of the comparative assessment of language policy options. Different policy choices are more or less conducive to economic prosperity, and we have seen that policies embracing diversity tend to be preferable in *contingent* terms, and probably in *absolute* terms as well. Just as economics is not merely concerned with the allocation of material and symbolic resources, but also with their distribution among individuals and groups, public policy is not just about efficiency, but also is about fairness; this also applies to language policies.

This raises another set of complex questions, which for lack of space can only be succinctly sketched here. Current research into these questions usually follows one of the two following approaches. On the one side, we find policy analysis, which is very closely connected to economics. Its application to language policies focuses on the identification and measurement of gains and losses that accrue, under different language regimes, to different groups (Gazzola 2016; Gazzola and Grin 2013). On the other side, normative political theory examines the ethical dimensions inherent to the comparison of those language regimes (De Schutter 2017; Shorten 2016). Both approaches largely concur that linguistic hegemony (independent of which language ends up in a dominant position), as well as processes of linguistic uniformization that lead up to it and are abetted by it, often prove regressive in socioeconomic terms and generate unjustifiable differences between language groups.

In other words, when language contact gives rise to unfair distribution of material and symbolic resources, economically inspired policy analysis dovetails with political theory to recommend redressing the situation through well-designed language policies.

Summing up, language economics and its application to language policies make critical contributions to our understanding of numerous aspects of the phenomenon of language contact. Research generally shows that, whether from an allocative or a distributive standpoint, linguistic diversity is more effective and fair than linguistic uniformity. This implies that linguistic diversity and the resulting language contact, if embraced and

properly managed, is economically beneficial and is more likely to deliver a fair society than a drift towards uniformity would. This suggests that public policies at the local, national, and inter/supranational levels should not only encourage individuals to develop a good portfolio of language skills, but also ensure that the world of business, science, and culture remains highly linguistically diverse.

6. References

Arcand, Jean-Louis & François Grin
2013 Language in economic development: Is English special and is linguistic fragmentation bad? In Elizabeth J. Erling & Philip Seargeant (eds.), *English and development: Policy, pedagogy and globalization*, 243–266. Bristol: Multilingual Matters.

Borooah, Vani, Donal Dineen & Nicola Lynch
2009 *Language and occupational status: Linguistic elitism in the Irish labour market*. Dublin: Economic and Social Research Institute.

Breton, Albert (ed.)
1999 *New Canadian perspectives: Exploring the economics of language*. Ottawa: Canadian Heritage.

Carrère, Céline (ed.)
2016 *L'impact économique des langues*. Paris: Economica.

Chiswick, Barry R. & Paul W. Miller
2007 *Economics of language: International analyses*. New York: Routledge.

Chiswick, Barry R., Harry A. Patrinos & Michael E. Hurst
2007 Indigenous language skills and the labor market in a developing economy: Bolivia. In Barry R. Chiswick & Paul W. Miller (eds.), *The economics of language: International analyses*, 473–489. London: Routledge.

De Schutter, Helder
2017 Two principles of equal language recognition. *Critical Review of International Social and Political Philosophy* 20(1). 75–87.

Esser, Hartmut
2006 *Migration, language and integration* (AKI Research Review 4; Arbeitsstelle interkulturelle Konflikte und gesellschaftliche Integration). Berlin: Wissenschaftszentrum Berlin für Sozialforschung.

Fürst, Guillaume & François Grin
2017 Multilingualism and creativity: A multivariate approach. *Journal of Multilingual and Multicultural Development*. http://www.tandfonline.com/doi/full/10.1080/01434632.2017.1389948. (accessed on 19 September, 2018).

Gazzola, Michele
2014 *The evaluation of language regimes: Theory and application to multilingual patent organisations*. Amsterdam & Philadelphia: John Benjamins.

Gazzola, Michele
2016 Multilingual communication for whom? Language policy and fairness in the European Union. *European Union Politics* 17. 546–569.

Gazzola, Michele & Bengt-Arne Wickström (eds.)
2016 *The economics of language policy*. Cambridge, MA: MIT Press.

Gazzola, Michele & François Grin
2013 Is ELF more efficient and fair than translation? An evaluation of the EU's multilingual regime. *International Journal of Applied Linguistics* 23. 93–107.

Gazzola, Michele & François Grin
2017 Comparative language policy and evaluation: Criteria, indicators and implications for translation policy. In Gabriel González Núñez & Reine Meylaerts (eds.), *Translation and public policy: Interdisciplinary perspectives and case studies*, 83–116. London: Routledge.
Gazzola, Michele, François Grin & Bengt-Arne Wickström
2016 A concise bibliography of language economics. In Michele Gazzola & Bengt-Arne Wickström (eds.), *The economics of language policy*, 53–92. Cambridge, MA: MIT Press.
Ginsburgh, Victor & Shlomo Weber (eds.)
2016 *The Palgrave handbook of economics and language*. London: Palgrave Macmillan.
Grin, François
1999 *Compétences et récompenses. La valeur des langues en Suisse*. Fribourg: Éditions Universitaires.
Grin, François
2003 Economics and language planning. *Current Issues in Language Planning* 4(1). 1–66.
Grin, François
2016 50 Years of economics in language policy. In Michele Gazzola & Bengt-Arne Wickström (eds.), *The economics of language policy*, 21–52. Cambridge, MA: MIT Press.
Grin, François & François Vaillancourt
2015 The economics of language policy: An introduction to evaluation work. In Francis M. Hult & David Cassels Johnson (eds.), *Research methods in language policy and planning: A practical guide*, 118–129. Boston: Wiley-Blackwell.
Grin, François, Claudio Sfreddo & François Vaillancourt
2010 *The economics of the multilingual workplace*. New York: Routledge.
Robbins, Lionel
1935 *An essay on the nature and significance of economic science*, 2nd edn. London: Macmillan.
Shorten, Andrew
2016 Four conceptions of linguistic disadvantage. *Journal of Multilingual and Multicultural Development* 38(7). 607–621.
Templin, Torsten, Andrea Seidl, Bengt-Arne Wickström & Gustav Feichtinger
2016 Optimal language policy for the preservation of a minority language. *Mathematical Social Sciences* 81. 8–21.
Vaillancourt, François (ed.)
1985 *Économie et langue*. Québec: Conseil de la langue française.
Zhang, Weiguo & Gilles Grenier
2013 How can language be linked to economics? A survey of two strands of research. *Language Problems and Language Planning* 37(3). 203–226.

François Grin, Geneva (Switzerland)

58. Educational sciences

1. Introduction
2. Language contact and education
3. Language policy, educational science, and multilingualism
4. Modernist assumptions in language contact and educational science
5. An alternative to traditional language contact study and education
6. Conclusion
7. References

1. Introduction

Scholars have long known and deplored the fact that schools tend to validate only standardized features of dominant languages and to interpret language differences as errors. Less attention has been paid to linguistic behaviors stemming from contact between languages. Here too, and even more strongly, schooling stigmatizes the behavior as flaws that need to be remedied. Students whose language practices are evaluated as being the result of language contact tend to receive lower grades on exams that time and again, and almost everywhere, validate only monolingual linguistic features that neatly fit as 'belonging' to the named language in which the evaluative instruments are written (Baker and Wright 2017; García 2009).

The stigmatization of practices that reveal what is seen as contact between languages occurs not only in monolingual education, but also in bilingual or multilingual education. In educational settings where more than one language is used as a medium of instruction, practices that fail to maintain separation between the languages are also regularly stigmatized. No less than monolingual programs, bilingual ones bracket one language from the other, devaluing the language practices of many bilingual and multilingual students and insisting that they behave in each language as their respective monolingual counterparts do (Grosjean 1982; Heller 1999). The effect of these imposed, artificial language practices in bilingual education is the same as in monolingual education: the 'other language' feature in bilingual education is treated in parallel to the substandard feature in monolingual education, with parallel negative effects. For language minoritized communities the effect is even more pernicious: bilingual schools most often construct the bilingual and multilingual practices of these communities as invalid.

The ways in which linguistic features and practices resulting from language contact have been constructed as harmful in schools assures dominant groups of a superior education, as it excludes those who use languages differently than do members of the elite majority. To change the educational reality of children whose language practices are the result of bilingualism, we would need to question the concept of language contact as it has been handed down to us by Weinreich ([1953] 1974), Haugen (1953), and the scholars who followed them (e.g. Matras 2009; Muysken 1985; Sankoff 2001; Thomason and Kaufman 1988), and we would have to shift our focus of attention from the language(s) to the speaker(s).

This paper traces the history of how educational science has used the concept of language contact to promote the bilingualism of the elite and/or to minoritize bilingual students whose language practices are different from those of monolinguals. We explore

https://doi.org/10.1515/9783110435351-058

how education for elite white students has interpreted language contact differently than for black/brown minoritized students. Finally, we propose an alternative conceptualization of language contact that may transform the ways in which we evaluate the linguistic performances of bilingual minoritized students, and thus, their education.

2. Language contact and education

It was the publication of Uriel Weinreich's *Languages in Contact: Findings and Problems* (1974) and Einar Haugen's *The Norwegian Language in America: A Study of Bilingual Behavior* (1953) that consolidated understanding of what we now know as studies of language contact. Among the most important contributions that Weinreich and Haugen made was to point out that the study of language contact could not focus narrowly on descriptions of language structures, but rather had to include social and psychological aspects. Weinreich (1974: 4) stated: "[p]urely linguistic studies of languages in contact must be coordinated with extralinguistic studies on bilingualism and related phenomena". The description of the structural linguistic phenomena of bilinguals was linked to speakers' social condition(s), including, according to Weinreich, geography, indigenousness, ethnicity, religion, race, sex, age, social states, occupation, and rural or urban context.

The sociolinguistic emphasis of studies of languages in contact was an important improvement over earlier accounts of what was simply described as 'language mixing'. Yet perhaps it was the overreach of language contact study into social issues that eventually had the most pernicious effect. Scholars studying education, and especially those specializing in the education of bilingual minoritized students, took up language contact as a most important factor in their educational outcome. In educational science, only one concept in Weinreich's influential book took root. That was the concept of *interference*. Although Weinreich credits the Prague school in Europe for having first examined interference from a structural standpoint, it is Weinreich's conceptualization, now linked to sociocultural factors, which has been most influential in educational science.

Weinreich (1974: 1) defines interference as "deviations from the norms of either language which occur in the speech of bilinguals". Weinreich continues by saying that interference "implies the rearrangement of patterns that result from the introduction of foreign elements into the more highly structured domain of language". Interference for Weinreich referred to applying knowledge from one language to another, a rearrangement that could be either positive or negative. There were two types of interference: (1) the positive one, in which speakers know the language to which the utterance belongs, and therefore, those that do not belong can be identified as borrowed or transferred; and (2) the negative one, in which there is no transfer in the structural sense because speakers have 'interlingual identifications' and cannot identify the language to which the utterance belongs. Educational science did not call the first type of contact interference, but rather termed it 'transfer' and applied it mostly to the education of the elite for bilingualism. In contrast, the term 'interference' was reserved for the second type of contact, which was identified mostly in the utterances of colonized and racialized powerless bilinguals. In Second Language Acquisition Studies, interference also led to the concept of 'interlanguage', referring to the presence of features of a first language when learning a second language (Selinker 1972). As a result, the education of bilingual students became focused

precisely on eradicating this interference, and not necessarily on educating them for success (see also Chapter 34).

3. Language policy, educational science, and multilingualism

As more and more former colonies in Africa and Asia became independent, the study of language policy was given priority. Sociolinguists were concerned with the 'language problems' of these newly emerging nations, and with their 'modernization' (Fishman 1972; Rubin et al. 1977). Language heterogeneity was seen as a problem to their advancement. In those early days of language policy studies, its three prongs − corpus planning, status planning, and acquisition planning − had much to do with purging contact between languages, that is, with ensuring that the chosen official languages − whether former colonial languages or indigenous languages − be used without any interference from the many other voices within the languages.

Although scholars involved in language policy activities saw themselves as working on behalf of multilingual populations, the emphasis on avoiding language contact and preventing interference meant that only a monoglossic concept of language was accepted. Language was to avoid the heteroglossia that the Russian Bakhtin Circle, including Mikhail Bakhtin himself (1981), as well as Vološinov (1995), described. Language was *langue* in the Sausurrean sense, a *langue* free of *parole*, a static sedimentation "deposited at the bottom of a lake", to use Weinreich's (1974: 11) metaphor. The speech of bilinguals was like "sand carried by a stream" (Weinreich 1974: 11). The task of the sociolinguist and of the educator was to pull the sand in the stream downwards so that it was deposited at the bottom, making it disappear. Eradicating interference became a mechanism of control of minoritized speakers, a means of governmentality (Foucault 2008) to use language to silence their voices and keep them oppressed. Language in schools became a technique for what Foucault (1990: 140) calls 'biopower', i.e.: "ways of achieving the subjugation of bodies and the control of populations". It was the education system that was called upon to stop the linguistic flows in the multilingual stream of communication and to control the bodies of those who were considered inferior.

4. Modernist assumptions in language contact and educational science

1953 was not only a big year for language contact studies, as the year of the publication of both Weinreich and Haugen's work. This was also the year in which UNESCO published its expert report titled *The Use of Vernacular Languages in Education.* The UNESCO report (1953: 6) stated:

> We take it as axiomatic that every child of school age should attend school and that every illiterate should be made literate. We take it as axiomatic too, that the best medium for teaching is the mother tongue of the pupil. But with any hundreds of languages lacking a *written form* [emphasis added], or when they have one, any literature for the use of pupils, it is difficult at present for both these aims to be pursued.

Three issues are prominent in the report: (1) schooling should be available for all children, (2) children should be educated in their mother tongue, and (3) given the fact that many languages are unwritten, it is difficult to teach every child to read or to offer mother tongue instruction. Besides establishing as axiomatic that an appropriate education for children should be in their home language, the 1953 UNESCO report served as the excuse for the failure to school millions of children − their languages were unwritten and thus it was impossible to educate them. Language was the culprit, the mechanism to blame for the failure to educate all children equally and justly.

The main flaw of the UNESCO 1953 report had to do with equating language and the language of schooling with western European epistemologies. Language in education was associated only with the named language of an imagined monolingual country, and not with the actual language practices of people. Language in education was also related only to the written form, an instrument of literacy.

Around this time, other educational science and language contact studies were mired in modernist assumptions about the mechanisms of human and social development. Rooted in structural-functionalist theory, they called for citizens' transformation from a traditional to a modern disposition (Peet 1999) that used standardized languages which were not 'primitive'. Languages were only perceived as valuable when they were standardized and codified in ways which reflected a single national identity that schools could then use as a medium of education. For speakers, this meant that only those who spoke a standardized language which had been assigned to a western nation-state were said to have a 'modern disposition'. The colonized and indigenous peoples of Asia, Africa, the Pacific, and the Americas were then said to have only a 'traditional' disposition that was expressed in what was considered a 'primitive' language. It is with these structural-functionalist and modernization understandings of speakers and their languages that education scholars took up the cause of language minoritized children around the same time.

As some started to experiment with bilingual education for language minoritized children in the newly independent and constructed nations of Africa and Asia, the paradigms of traditional monolingual education did not change. Bilingual education was now expected to take place in two standardized languages instead of one. This reduced bilingual education to simply the pluralization of monolingual education, and educational achievement continued to be measured according to the pupils' performance in a standardized language, usually only the dominant one. Minoritized bilingual students whose language performances were seen as instances of 'interference' and language contact were then reduced to 'semilinguals', a term used by some applied linguists around this time.

4.1. Semilingualism

The concept of semilingualism gained traction in the 1960s and 1970s, especially in Sweden. Skutnabb-Kangas (1981: 250) attributes the term to a book by Hansegård. Semilingualism was used to refer to the poor competence of Finnish students in Sweden in both Finnish and Swedish, a product of language contact produced by social factors.

The concept of semilingualism was derived from the idea of negative interference in situations of language contact. As a scholarly pursuit, work on 'semilingualism' had a short life, as scholars opposed its deficit view of children's language (Edelsky et al.

1983; Martin-Jones and Romaine 1986). And yet, as ways to educate bilingually were developed in the following decades, the specter of semilingualism led to language regimentation that ensured there would be no language contact.

4.2. Education for bilingualism: early efforts

Education for bilingualism is not new (Mackey 1978). From its beginning, however, it was the teaching of an additional written language to the elite as an enrichment activity that encouraged its growth. In Roman aristocratic circles, for example, Greek-Latin bilingual education was a way to learn the written language of the admired Hellenic civilization. The early focus of bilingual instruction was not on teaching students to speak the language, but rather merely on reading it because of its classical or sacred value. There was no conflict (or contact) between the language practices of the home and community, and those of the classical and sacred texts studied in schools. There was no question about which language an utterance belonged to. Nonbelonging, when it did occur, could "be separated as borrowed or transferred" (Weinreich 1974: 7).

It was precisely that tradition of teaching a new additional language that was the basis for foreign language teaching. For example, when the Modern Language Association was established in the United States in 1883, it established as its goal the study of "literacy culture, philological scholarship and linguistic discipline" (as cited in García [1993: 75]). And in 1929, the US Modern Foreign Language Study recommended reading as the primary aim of foreign language study. The two languages were clearly demarcated − one spoken, the other in the printed text; one spoken by all, the other written by the cultured class and the object of linguistic study. Language contact phenomena were then not frequently attested, and when they did occur, it was easy to determine which utterances did not belong and to identify them as borrowed or transferred.

Yet as education took on its modernization role, the goal of language education, as well as its pedagogical strategies, started to shift, as more and more learners became involved in schooling for bilingualism and as more and more bilingual minoritized students clamored for social justice and an equal education.

4.3. Bilingual teaching: teaching an additional language or teaching bilingual children

There is an important difference between teaching language majority youth an additional language (education for bilingualism), and teaching bilingual minoritized youth (educating with bilingualism). We expand below on the types of education programs that have been developed for both purposes, while considering the ways in which scholarship on language contact has impacted education programs.

4.3.1. Teaching an additional language to language majorities: education for bilingualism

Teaching language majority children an additional language usually occurs *sequentially* (that is, after the first language is acquired) and for enrichment purposes. The type of

bilingualism that is produced is considered a resource no matter its characteristics, and is referred in the literature as *elite bilingualism* (Fishman 1977).

Foreign language education programs have traditionally been the most popular way of teaching students an additional language. These programs are prevalent in middle and secondary schools, and usually start after the children have developed literacy in what is considered their first language. There are limited expectations that the students will become bilingual, and language contact is seen as simply transfer of features that can be easily identified as belonging to the other language.

Content and Language Integrated Learning (CLIL) programs have been developed to address the failure of foreign language programs to develop bilingualism. As in bilingual education programs, CLIL programs use the target language as a medium of instruction.

Immersion bilingual education programs have also been developed for language majority families anxious to ensure that their children become bilingual. In these programs, the target language is used as medium of instruction almost exclusively at the beginning of the program (total immersion) or at least 50 % (partial immersion), and they can start at the beginning of formal education or later. In immersion bilingual education, the languages are used as media of instruction and are also clearly separated.

Foreign language, CLIL, and immersion bilingual education programs display what we might call a conflicted ideology: on the one hand, they defend a strict monoglossic understanding of bilingualism, but on the other hand, they accept language contact phenomena as natural borrowings or transfers. The goal is to be protective of the first dominant language and tolerant of language contact phenomena, which are considered positive transfers from the first language to the second language. Language contact for these majority students is always considered positive, and all of their bilingual performances are considered resources (Ruiz 1984) for learning, despite the monoglossic arrangements of the education programs.

4.3.2. Teaching bilingual minoritized youth: educating with bilingualism

In contrast, the bilingualism of indigenous, conquered, colonized, refugee, and/or immigrant youth in schools (what Fishman [1977] called *folk bilingualism*) is often perceived as a problem. The designation of this bilingualism as ‘folk’ tells a story of being stigmatized as informal, as well as traditional or primitive, and not in any way modern. Folk bilingualism is connected to Weinreich’s perception that these bilingual speakers have ‘interlingual identifications’ which increase the overlap between the two languages to the point that one cannot speak about transfer from one language to another. Instead, these speakers and their listeners are said to often not know to which language the utterance belongs. It is this type of interference that is always evaluated negatively.

It is no surprise then that education programs to educate language minoritized speakers, whether monolingually or bilingually, have always operated in ways that minimize what Weinreich (1974: 7) calls “interference that does not involve transfer”. In other words, educational programs for language minoritized bilinguals often focus on differentiating between the two languages. To do so, education programs to educate bilingual youth of all types tend to strictly separate languages, that is, they follow a diglossic approach (Fishman 1965) in which the two languages are strictly compartmentalized through different teachers, subjects, times of day, physical spaces, etc.

We consider here the four most prevalent types of educational programs for minoritized bilingual students: (1) monolingual instruction in a second/additional/new language; (2) transitional bilingual education; (3) developmental bilingual education; and (4) immersion revitalization bilingual education. As we will see, regardless of the type of educational program these bilingual minoritized children receive, the goal of education is clearly to eradicate language contact phenomena considered to be negative interference. The expectation is that the language performances of minoritized bilingual youth be eradicated of their fluidity, of the 'sand in the stream'. The goal is for these bilingual youth to become either monolingual or be able to perform 'dually' as two monolinguals in one. As we will see, in both cases, these minoritized bilingual youth will not be able to perform as monolinguals in either one language or two. The normalization of this evaluative scheme in schools results in what we have been taught to call 'the achievement gap', locating the failure in the students and their communities rather than facing our failure to understand the dynamic bilingualism (García 2009) at play.

Monolingual second language programs limit the language use of the bilingual learner to the target second/additional/new language. Language contact is avoided. *Transitional bilingual education programs* gradually erase the language of the learner, substituting it with the dominant language of the society. The goal of second language programs as well as transitional bilingual education programs is to ensure that all types of dynamic bilingual performances are eradicated.

Developmental bilingual education programs (labeled *dual language education programs* in the United States) aim to develop the dominant language of the society and an additional language (usually a minoritized language of bilingual communities that reside locally). *Immersion revitalization bilingual education programs* have as their goal the reversal of the language shift of indigenous and colonized groups whose language practices had been mostly eradicated by a monolingual school system. And yet, the goal of both developmental bilingual education and immersion revitalization bilingual education is to ensure that the bilingual performances of youth are 'dual', that is, that language contact of all types is avoided. The thinking is that education must counteract what is seen as the negative interference that occurs when minoritized speakers use their language. The goal of education must be to make these bilingual speakers strictly separate their language practices so that there is no contamination between one language and the other. The downside of this 'dual' approach is that the emphasis of education becomes the avoidance of language contact, rather than the bilingual education of the child.

5. An alternative to traditional language contact study and education

The concept of language contact has been based on the natural − as opposed to the socially invented − existence of two or more named languages. It assumes that bilingual/multilingual speakers have two or more language systems, and that when features from system A come into those of system B (or vice versa), the languages are contaminated. In the case of elite bilinguals and the education programs that have been developed for them, that 'interference' is sometimes considered 'transfer', and viewed as a resource. In the case of minoritized bilinguals and the education programs that have been devel-

oped for them, that 'interference' is mostly seen as a problem, as negative, and as something that has to be avoided. Relying on this last (mis)conception, bilingual schools for minoritized bilinguals maintain the social and linguistic order by demanding that bilingual students always separate their languages. To do so, schools demand that only one language be used at one time, and that it be used without any trace of the other one. Under this guiding principle, bilingual students from language minoritized communities can never be equal to those from language majority communities who can use the dominant language without any trace of language contact.

An alternative language contact conceptualization relies on the understanding of the bilingual speakers' language practices from the point of view of the speakers themselves, and not from the perspective of the two named languages. From the perspective of bilingual speakers, their linguistic repertoire does not consist of two named languages, but of a unitary system of linguistic features that they use dynamically and simultaneously to make meaning. We, along with other scholars, have used the term 'translanguaging' to name this theoretical position (García and Li Wei 2014; Li Wei 2018; Otheguy, García, and Reid 2015, 2018).

In many ways, translanguaging theory expands what Weinreich, writing about language contact from the external perspective of named languages and dialects, termed 'interlingual identifications'. For us, however, the identifications are *intra*lingual, that is, occurring within the internal unitary language system of the bilingual speaker, and not across two socially constructed named languages. Although we do not in any way dismiss the sociopolitical reality of named languages, and its consequences on the lives of minoritized bilinguals, we wish to dispose of the concept of interference in the ways described by Weinreich, whether in positive or negative ways.

Jim Cummins (1979) has been one of the most influential scholars in studies of bilingualism and bilingual education. His theories of linguistic interdependence and of linguistic transfer from L1 to L2 have been helpful in legitimizing bilingual education. Translanguaging theory goes *beyond* Cummins' notion of transfer because it goes *beyond* language. Instead, translanguaging posits that named languages have social reality, but no linguistic reality (Otheguy, García, and Reid 2015, 2018). One's own language, as Bakhtin (1981: 66) has said, "is never a single language; in it there are always survivals of the past and a potential for other-languagedness". Named languages have been inventions (Makoni and Pennycook 2007: 14) that were "administratively assigned to colonized populations as mother tongues". And education for minoritized populations that simply uses these inventions, while excluding all other practices that do not fall neatly within the norms and conventions of what has been constructed as the school language(s), is a mechanism to further oppress colonialized and racialized groups. Their own bilingual language practices at home and communities are made invalid; meanwhile, the only way to acquire the invented named language, now designated as the group's mother tongue or vernacular, is through the formal education that in turn excludes the group.

Language contact theorists rightly engage with the social factors that create the contact. Yet their understandings of language have been shaped by structural-functional and modernist concerns about human and social development that have left out the microphysics of power (Foucault 1990). An understanding of how linguistic normalization and language invention have been a mechanism of power is most important in shifting

our understanding of language contact and clarifying the concept of translanguaging (García and Li Wei 2014; Otheguy, García, and Reid 2015).

If the language practices of bilinguals are seen through the lens of negative interference, governable subject positions or 'docile bodies' can be produced (Foucault 1979). The alternative view that we present here in taking up translanguaging opens up spaces for the legitimization of new subjectivities, speaking and listening subjects whose fluid language practices are not evaluated as negative interference. As such, bilingual youth translanguaging can be leveraged in their education (e.g. Creese and Blackledge 2010; García, Johnson, and Seltzer 2017; García and Kleyn 2016; Gort 2015; Hornberger and Link 2012; Palmer et al. 2014) in both instruction and assessment. Our conceptions of language cannot be separated from the sociohistorical, political, and economic conditions that produce them.

Unlike the structural-functionalist beginnings of language contact theory and its modernist applications in education, translanguaging resides in a space that Mignolo (2000: xxv) describes as going beyond "the colonial difference". In this translanguaging space (Li Wei 2011), Mignolo (2000: xxv) explains that "the restitution of subaltern knowledge is taking place [and] border thinking is emerging". This translanguaging space constitutes a transformation of the modern/colonial world system and a decolonization of Western epistemologies. The rigidity of epistemic frontiers, explains Mignolo (2000: 12), were "established and controlled by the coloniality of power in the process of building the modern/colonial world system". Named languages are legacies of European colonialism, and of the practices and forms of knowledge of European colonialism. As such, they function in education as instruments of what Aníbal Quijano (2000) calls 'coloniality of power'. Named languages are linked to the racial, political, and social hierarchical orders imposed by European colonialism that privilege some, while marginalizing others.

Translanguaging in the education of bilingual minoritized speakers acknowledges that all the features of their repertoire are always present. It is not that there is contact across languages, it is that language is, as Bakhtin (1981: 291) has said, "heteroglot from top to bottom". The task of educators must be to leverage bilinguals' entire language repertoire in order to meaningfully and equitably educate students whose language features will always differ from those of dominant monolingual speakers.

Because named languages are important in schools and society, the planning of language use is important for instruction. However, equally as important is to acknowledge translanguaging as a resource, and not to see it as simply language contact, or as a deviation from monolingual elite norms. To truly educate all students, educators need to use their entire linguistic system, not temporarily as a simple scaffold, but as a way to transform how we educate.

6. Conclusion

If educators would conceptualize the bilingual student not in terms of language contact and interference, which puts languages in different, separate spaces, but instead in terms of a continuum of language practices that bilinguals and others have, bilingual youth linguistic practices would be evaluated differently. Educators would then not simply evaluate the use of the conventional language features of each language. They would

instead focus on how students use their language repertoire to gather, comprehend, evaluate, synthesize, and report on information and ideas, to engage with complex texts, and to write to persuade, explain, and convey experiences. Only when educational science frees itself from improbable conceptualizations of language contact and traditional bilingual behavior will we be able to educate all our citizens meaningfully.

It is in schools where translanguaging meets with more resistance. In homes and communities, translanguaging is a useful resource to communicate with many. The focus of educational science cannot be the eradication of the effects of what is seen as language contact. Education must engage students, and especially minoritized bilingual youth, with their own *translanguaging power* to make meaning, learn, and overturn the biopower of schools that subjugates their bodies and tongues.

Studies of translanguaging in the education of bilinguals are growing, although the practice cannot be said to be mainstream. Its study has the potential to transform schools and classrooms, as well as the future of language minoritized children. In addition, however, the close study of translanguaging in classrooms has the potential to bring the classical scholarship on language contact up to date with critical poststructuralist understandings of language in society.

Acknowledgments

I'm grateful to Ricardo Otheguy for his reading of a first draft of this article.

7. References

Baker, Colin & Wayne Wright
2017 *Foundations of bilingual education and bilingualism*, 5th edn. Bristol, UK: Multilingual Matters.

Bakhtin, Mikhail
1981 *The dialogic imagination.* Austin: University of Texas Press.

Creese, Angela & Adrian Blackledge
2010 Translanguaging in the bilingual classroom: A pedagogy for learning and teaching? *The Modern Language Journal* 94(1). 103–115.

Cummins, Jim
1979 Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research* 49(2). 222–251.

Edelsky, Carol, Sarah Hudelson, Barbara Flores, Florence Barkin, Bess Altwerger & Kristina Jilbert
1983 Semilingualism and language deficit. *Applied Linguistics* 4(1). 1–22.

Fishman, Joshua A.
1965 Who speaks what language to whom and when? *La Linguistique* 2. 67–88.

Fishman, Joshua A.
1972 *Language in sociocultural change: Essays by Joshua A. Fishman.* Stanford, CA: Stanford University Press.

Fishman, Joshua A.
1977 *Bilingual education: Current perspectives: Volume 1: Social Science*. Arlington: Center for Applied Linguistics.

Foucault, Michel
1979 *Discipline and punish: The birth of the prison*. New York: Vintage Books.
Foucault, Michel
1990 *The History of Sexuality: Volume 1: An introduction*. [1978 English translation by Random House]. New York: Vintage Books Edition.
Foucault, Michel
2008 *The birth of biopolitics: Lectures at the College de France, 1978−79.* New York: Palgrave MacMillan.
García, Ofelia
1993 From Goya portraits to Goya beans: Elite traditions and popular streams in US Spanish language policy. Southwest Journal of Linguistics 12(1−2). 69−86.
García, Ofelia
2009 *Bilingual education in the 21st century: A global perspective*. Malden, MA & Oxford, UK: Wiley/Blackwell.
García, Ofelia & Li Wei
2014 *Translanguaging: Language, bilingualism and education*. London, UK: Palgrave Macmillan Pivot.
García, Ofelia & Tatyana Kleyn (eds.)
2016 *Translanguaging with multilingual students*. New York and London: Routledge.
García, Ofelia, Susana Johnson & Kate Seltzer
2017 *The translanguaging classroom: Leveraging student bilingualism for learning.* Philadelphia: Caslon.
Gort, Mileidis
2015 Transforming literacy learning and teaching through translanguaging and other typical practices associated with "doing being bilingual". *International Multilingual Research Journal* 9(1). 1−6.
Grosjean, François
1982 *Life with two languages.* Cambridge, MA: Harvard University Press.
Haugen, Einar
1953 *The Norwegian language in America: The Bilingual community.* Philadephia: University of Pennsylvania Press.
Heller, Monica
1999 *Linguistic minorities and modernity: A sociolinguistic ethnography.* London: Longman.
Hornberger, Nancy H. & Holly Link
2012 Translanguaging and transnational literacies in multilingual classrooms: A bilingual lens. *International Journal of Bilingual Education and Bilingualism* 15(3). 261−278.
Li Wei
2011 Moment analysis and translanguaging space: Discursive construction of identities by multilingual Chinese youth in Britain. *Journal of Pragmatics* 43(5). 1222−1235.
Li Wei
2018 Translanguaging as a practical theory of language. *Applied Linguistics* 39(1). 9−30.
Mackey, William
1978 The imposition of bilingual education models. In Jim Alatis (ed.), *Georgetown University round table on languages and linguistics*, 1–18. Washington, D.C.: Georgetown University Press.
Makoni, Sinfree & Alistair Pennycook
2007 *Disinventing and reconstituting languages.* Clevedon, UK: Multilingual Matters.
Martin-Jones, Marilyn & Suzanne Romaine
1986 Semilingualism: A half-baked theory of communicative competence. *Applied Linguistics* 7(1). 26−38.
Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.

Mignolo, Walter
2000 *Local histories / global designs: Coloniality, subaltern knowledges, and border thinking.* Princeton: Princeton University Press.
Muysken, Pieter
1985 Linguistic dimensions of language contact: The state of the art in interlinguistics. *Revue Québécoise de Linguistique* 14. 49–76.
Otheguy, Ricardo, Ofelia García & Wallis Reid
2015 Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review* 6(3). 281–307.
Otheguy, Ricardo, Ofelia García & Wallis Reid
2018 A translanguaging view of the linguistic system of bilinguals. *Applied Linguistics Review.* https://doi.org/10.1515/applirev-2018-0020.
Palmer, Deborah K., Ramón A. Martínez, Suzanne G. Mateus & Kathryn Henderson
2014 Reframing the debate on language separation: Toward a vision for translanguaging pedagogies in the dual language classroom. *The Modern Language Journal*. 98(3). 757–772.
Peet, Richard
1999 *Theories of development.* New York: Guilford Press.
Quijano, Aníbal
2000 Coloniality of power, Eurocentrism, and Latin America. *Nepantla: Views from South* 1(3). 533–580.
Rubin, Joan, Björn H. Jernudd, Jyotirindra Das Gupta, Joshua A. Fishman & Charles A. Ferguson (eds.)
1977 *Language planning processes*. The Hague: Mouton Publishers.
Ruiz, Richard
1984 Orientations in language planning. *NABE Journal* 8(2). 15–34.
Sankoff, Gillian
2001 Linguistic outcomes of language contact. In Peter Trudgill, Jack Chambers & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 638–668. Oxford: Basil Blackwell.
Selinker, Larry
1972 Interlanguage. *International Review of Applied Linguistics* 10(1–4). 11–23.
Skutnabb-Kangas, Tove
1981 *Bilingualism or not? The education of minorities.* Clevedon, Avon, UK: Multilingual Matters.
Thomason, Sarah & Terrence Kaufman
1988 *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
UNESCO
1953 *The use of vernacular languages in education*. Paris: UNESCO.
Vološinov, Valentin. N.
1995 Marxism and the philosophy of language [1929]. In Pam Morris (ed.), *The Bakhtin reader: Selected writings of Bakhtin, Medvedev and Vološinov*, 26–37. New York: Oxford University Press.
Weinreich, Uriel
1974 [1953] *Languages in contact: Findings and problems*. The Hague: Mouton.

Ofelia García, New York, NY (USA)

59. Language contact and legal studies

1. Introduction
2. Law and language: a complex relationship
3. Legal regulations of languages in contact: trends and debates
4. Language contact, law, and justice
5. References

1. Introduction

Monolingual contexts do not necessarily require legal responses; nor do multilingual contexts. However, given the fact that language contact in an increasingly multilingual world often goes hand in hand with latent (or more manifest) forms of language conflict, legal responses to prevent or neutralize language conflict are often required. A legal studies approach to language contact considers language to be a relevant communicative instrument that is the object of both regulation and research. The legal studies approach to linguistic phenomena is a complex one, in that it is fraught with issues that are heavily debated in academia as well as in politics. While that part of legal studies which concentrates on Language Law is preoccupied with a more objective and normative approach to social-linguistic phenomena, the linguistic (human) rights approach is more subjective, in that it necessarily concentrates on the existence and nature of the rights of (groups of) speakers. After a brief sketch of the complex relationship between 'language' and 'Law' (Section 2), this chapter will concentrate on 'Language Law' (Section 2.1) and 'linguistic (human) rights' (Section 2.2) as important complementary and interacting fields of legal studies. It then turns to a discussion of current trends and debates on legal regulations of language contact (Section 3), before briefly discussing language contact, law, and justice (Section 4).

2. Law and language: a complex relationship

Language has been given attention by Law primarily due to the fact that it is an instrument of social relations and power. Law obviously also needs language, since it is produced and implemented through language(s), and as such, has a significant impact on the symbolic as well as political value of a given language (variety) in a given setting.

Explicit legal regulations concerning the use of languages can be found as early as the Middle Ages, but even before that, customary rules governed the use of languages in various domains of language use. In modern times, the relationship between language and Law became more relevant and prominent. With the consolidation of the modern state, lawyers started paying more attention to language realities. In the age of nationalism, language came to be perceived as an essential element of the legal and political recognition of defined collective identities (Ruiz Vieytez 2001). In order to strengthen their internal cohesion as 'imagined communities' (Anderson 1983), the new independent states of the 19th and 20th centuries as well as the political movements around national minorities produced legal proposals concerning their respective languages.

https://doi.org/10.1515/9783110435351-059

The study of the complex interplay between language and Law was mainly advanced in research dealing with legal regulations on languages in multilingual states such as Belgium (1873), Canada (1867), Finland (1922), Ireland (1937), or South Africa (1910), which were among the first to regulate official uses of more than one language. In addition, the fate of linguistic minorities in the Interbellum also contributed to the attention that was given to language contact in legal studies (Balogh 1930; Fouques Duparc 1922). The more systematic study of normative Language Law and linguistic (human) rights only really started in the 1980s and 1990s with the work of Skutnabb-Kangas (2000), Skutnabb-Kangas and Phillipson (1995), Turi (1994), De Witte (1985), Kontra (1999), De Varennes (1996), Woehrling (2005), and others. More recently, academic debates have broken new ground and started to concentrate on the concept of Linguistic Justice, starting with the seminal work on this topic by Philippe Van Parijs (2011). In this debate, other scholars, like Mowbray (2012) and De Schutter (2007, 2016), have made significant contributions, but the presence of legal scholars in the field is still rather scarce.

Studying the complex relationship between language and Law is quite a challenge for at least two reasons:

1. The regulation of languages by means of legal norms is complicated by the fact that multilayered concepts − such as 'language' and 'dialect' − that include references to languages as codes of communication, as well as to languages as context-dependent cultural products or components of a collective identity, are not easily transferrable into universally applicable legal categories (Kontra 1999). The same legal and political concepts convey different meanings in different countries, and on top of that, no standardized terminology exists (Henrard 2003: 40). That holds for 'official language', 'state language', and other concepts that appear in constitutional regulations, such as 'minority language', 'regional language', 'language of interethnic communication', 'indigenous' or 'native language', 'language of preference', and other similar categories (Ruiz Vieytez 2005; Turi 2015).
2. The direct impact of legal regulations on linguistic reality or the course of linguistic processes is uneven and cannot be taken for granted. As argued by Hogan-Brun and Wolff (2003: 5), social and economic factors sometimes have a far stronger immediate impact on language dynamics than does the adoption of legal regulations. Even coercion, the principal defining characteristic of legal rules, has had limited effects on the dynamics of language contact. Legal sanctions are often less efficient than social sanctions, such as the loss of social prestige (Viaut 1999), difficulty in conducting a profitable business, or school failure. However, legal regulations also have both a symbolic effect and a real impact with regard to the importance given to a particular language by the community (Willemyns 1992: 3−16). The declaration of a minority language as official helps in raising its social prestige and modifying sociolinguistic perception (see e.g. the case of Basque and Catalan in Spain). The legal recognition of the official status of a language is of major importance in language maintenance and shift, although it is never definitive (Romaine 2002: 2). Apart from symbolically 'shaping' social reality by highlighting the presence, importance, or absence of specific language communities, legislation is a highly effective means to attribute power and influence to one language (variety), while simultaneously obscuring or discriminating against other varieties. Through the aforementioned and related categories, Law thus

also helps to hide, create, or transform linguistic realities, for instance, by promoting some varieties of a given language over others (May 2012: 160).

The above points have to be taken into account in the two different major strands of legal studies that deal with languages, namely: (1) the strand that concentrates on Language Law as a specific branch of the legal order, analyzing language regulations as a normative system from a global (objective) approach; and (2) the strand that focuses on linguistic (human) rights, analyzing language regulations from the (subjective) point of view of citizens and their rights.

2.1. Languages from the state perspective: language law

When studying Language Law from a state perspective, it is essential to consider its multilevel dimension. Legal regulations of language are not only to be found at the state level. In certain countries, autonomous regional or local bodies may have considerable jurisdictional capacity to regulate the recognition and use of languages, as well as specific linguistic rights, within their territories (examples include Catalonia in Spain, Greenland in Denmark, the Aland Islands in Finland, autonomous republics of Russia, Hong Kong in China, New Brunswick in Canada, municipalities in Switzerland, and so on). At the international level, different organizations have their own linguistic regimes, and there are a number of international treaties that – either entirely or partly – deal with languages and linguistic rights, and to a certain extent, even interfere with other regulations at the level of the states belonging to these associations (for instance, the *European Charter for Regional or Minority Languages* or the *Convention on the Rights of the Child*). Legal provision on languages can thus respond to a large variety of sources.

From a legal objective approach, Language Law can be defined as the set of legal rules and principles governing the use of languages in a given political community (Guillorel and Koubi 1999: 22; Vernet 2003: 22). Language Law includes both the major legal and political decisions concerning languages, and the main rules that articulate the language policies of each political community. Besides the institutionalization of official languages, Language Law may also include clauses prohibiting linguistic discrimination, recognition of specific linguistic rights, rules for the institutional use of languages, rules concerning the linguistic corpus of a specific language, promotion and protection clauses, rules concerning the distribution of powers, and rules about language requirements for access to specific functions. Yet Language Law may still be challenged for its lack of systematicity, and for the (sometimes high) level of legal insecurity this engenders for the population (Fernández Liesa 1999: 14).

The state-boundedness of Language Law as a discipline notwithstanding, a comparative study of Language Law has come into existence that has gained traction since the 1990s (De Varennes 1996; Milian i Massana 1994; Pizzorusso 1993; Turi 1994; Vernet 2003; De Witte 1985). One of the outcomes of the comparison of legal provisions that establish the official status of the languages spoken in different states all over the globe is a typology of legal linguistic models, which can be summarized as follows:

a) Official multilingualism: The state recognizes more than one official language for the entire territory of the country. That is the case in Belarus, Bolivia, Cyprus, Fiji, Finland, Ireland, Israel, Malta, Paraguay, Singapore, and South Africa.

b) Territorialized official multilingualism: There are several official languages, but these are official only in certain areas of the state territory (with the exception of some bilingual/multilingual concrete regions or institutions). Examples include Belgium, Bosnia-Herzegovina, Canada, and Switzerland.
c) Asymmetric official multilingualism: The state has one official language for the entire country, but in some politically autonomous areas, there is a second (or third) official language. Under this model, a part of the country has a single official language, while in other areas, some minority languages are given co-official status. This model is in vigor in China, Denmark, India, Italy, Moldova, Russia, and Spain.
d) Limited official monolingualism: There is a single official language for the entire state, but some minority languages are protected by means of special measures, or have an official status only at the local (administrative) level. Examples include Austria, Bulgaria, Croatia, Germany, Hungary, Norway, Portugal, Serbia, Slovakia, Slovenia, and The Netherlands.
e) Official monolingualism: There is a single official language, and no minority language is explicitly protected at the constitutional level. This model does not exclude cases in which the state has ratified some international regulations that will ultimately recognize linguistic rights, but no official status has been formally given to any other languages. Examples include Australia, Greece, Iceland, Japan, Oman, and Turkey.

Obviously, these models are not absolute, in the sense that only one model would apply to a given state. A country like Italy, for example, in fact falls under models (c), (d), and (e) with regard to different minority languages (i.e. to German, Slovene, and Catalan, respectively), as does the Russian Federation. In addition, the distance between the formal configuration (expressed in the Law) and the material reality (expressed in sociolinguistic data) of the different models should also be considered, since gaps can be observed in any context, and therefore in any of the aforementioned models.

When moving from the domestic legal framework to International Law, it becomes more difficult to speak of the existence of a consistent Language Law. In principle, states are free to decide which languages will be officially used in their territory, and International Law may only establish certain limits to state power through the protection of human rights. In addition to clauses prohibiting discrimination on linguistic grounds, certain rights of persons belonging to linguistic minorities have been included (Article 27 of the *International Covenant on Civil and Political Rights* and Article 30 of the *Convention on the Rights of the Child*). Concerning the languages of indigenous peoples, the most relevant international legal document is the *ILO Convention No. 169*. In Europe, both Article 22 of the *Charter of Fundamental Rights of the European Union* and the *European Charter for Regional or Minority Languages* (ECRML), currently the most thorough piece of legislation on languages (Grin 2003; Nogueira, Ruiz Vieytez, and Urrutia Libarona 2012; Woehrling 2005) should be mentioned. The *Framework Convention for the Protection of National Minorities* is also relevant in this regard. In addition, some soft-law international documents establish principles in the field, like the *Oslo Recommendations Regarding the Linguistic Rights of National Minorities*, the *UN Declaration of the Rights of Persons Belonging to National or Ethnic, Linguistic or Religious Minorities*, and the *UN Declaration on Indigenous Peoples*. Other international documents adopted by nongovernmental actors include the *Universal Declaration of Linguistic Rights* and the most recent *Protocol to Ensure Language Rights*, adopted in 2016.

To sum up, the number and rank of the rules governing the use of languages have increased significantly in recent decades. On the one hand, this is due to the fact that language contact has grown as a result of population movements and the transformation of communications. On the other hand, the various political and social actors seem to increasingly recognize the crucial role played by languages in the organization of society. However, the weakness of international regulations and the high (normally constitutional) rank of domestic Language Law constrain our ability to take a more inclusive and democratic legal approach to linguistic diversity in our societies. This confirms the limitations of the Law when addressing contact between languages, since it always tends to reflect and maintain the power relations that occur *de facto* within each society.

2.2. Languages from the speakers' perspective: linguistic (human) rights

When the relationship between language contact and Law is considered from the point of view of speakers, the most relevant concept is that of linguistic rights, that is, the linguistic options that are legally protected either by domestic or international legal systems. The declaration of a language as 'official' automatically generates linguistic rights, but these can also be partially recognized by other means, as with the ratification of international treaties or domestic legislation containing specific rights for certain groups.

The concept of linguistic rights (or linguistic human rights; Skutnabb-Kangas and Phillipson 1995) as a separate category can be challenged. The expression 'linguistic rights' usually refers to the existence of clearly established rights for certain groups or territories, but very often, language is a necessary element for enjoying civil or political rights. In the first case, we would be talking about explicit or specific linguistic rights, which are rights expressly adopted to allow the use of certain languages (e.g. the right to address a public authority in one's own language). In the latter, we refer to implicit or generic linguistic rights, which are linguistic applications of civil or political fundamental rights (e.g. the right to choose the language for advertising as part of freedom of expression or the right to a fair trial that includes the guarantee of understanding the accusation).

Linguistic rights may thus refer to language in general, or to a specific language chosen by an individual or a group. The 'right to language' (Article 18 of the Swiss Constitution) is seen as a fundamental and universal human right (Turi 1990: 60) that legitimizes people to use their language in all private transactions and in certain public relations, irrespective of the official status of that language. This general right to language comprises a set of linguistic contents within human rights, such as freedom of expression, the right to a name, the right to private and family life, and freedom of education (De Varennes 2001). The 'right to a language', on the contrary, involves the power of a linguistic group to obtain a particular legal status for its own particular language which is embodied in other rights, such as the use of that language in public relations. This specific right to a language (or to linguistic differentiation) only becomes effective once it is incorporated into positive norms (Turi 1994: 116; Turi 2015: 14).

These linguistic rights can be relevant and implemented in both private and public domains. Depending on the corresponding legislation, they may be enjoyed according to the territoriality principle or the personality principle (Reaume 2003; Skutnabb-Kangas

and Phillipson 1995). The former implies that linguistic rights extend to all residents of a given territory, be they foreigners or nationals, insofar as they find themselves within the limits of that territory. The latter means that the possibility of using a particular language in different public or private domains applies to the members of a given group, regardless of their place of residence.

Other debates concerning the nature of linguistic rights refer to the holder of the rights. Most of the existing linguistic rights must be exerted by individuals, but in many cases, collective organizations can also claim them. Many linguistic rights can be defined as collective rights, in the sense that they protect a collective element of identity and communication. However, from a strictly legal perspective, the rights tend to be attributed to individual speakers, even if very often a minimum number of them is required to be put into practice (as happens with other individual rights, such as the right to marriage or the right to associate).

Linguistic rights usually derive from the legal recognition of a language's official status. But Law can also recognize explicit linguistic rights for certain minorities for political reasons. In some European countries, the implementation of the ECRML has served to develop protection standards for minority languages, although only benefiting the languages considered to be 'traditionally spoken' and not 'languages of immigrants'. Here, a constant relationship emerges between Law, language, and national identity (Ruiz Vieytez 2009). An important issue is the current debate about the extension of linguistic rights for speakers of traditional minority languages to the speakers of other, so-called immigrant languages. Many of these highly relevant debates remain unsolved, and show that the theory of linguistic rights is still in need of consolidation and further argumentation (Grin 2005: 451).

Language can also be used as a stepping stone for access to other rights. In addition to the legal restrictions imposed on the possible official use of nonnative languages, access to certain political positions or professions may be subject to linguistic constraints. This is evident in the processes involved in acquiring citizenship from States and obtaining residence permits. A trend in recent decades has been to impose on those who aspire to naturalization stronger, more formal requirements for knowledge of the State's official language (Extra, Spotti, and Van Avermaet 2009; Van Oers, Ersboll, and Kostakopoulou 2010; Ryan and Mcnamara 2011: 180–194).

Finally, it is worth mentioning that, in a number of countries, there are specific institutions endowed with the responsibility for guaranteeing linguistic rights. They normally consist of language boards or language commissioners that monitor the protection of certain language rights. This is the case of the Language Commissioners in Canada (Howard 2007: 4) and Wales (Mac Golla Chríost 2016), or the Pan South Africa Language Board (Beukes 2009: 48–50). Obviously, apart from these specific bodies, linguistic rights can also be protected through general legal mechanisms, such as judicial authorities, general ombudspeople, or administrative offices.

3. Legal regulations of languages in contact: trends and debates

In a global analysis of the existing regulations regarding languages, three different trends can currently be observed. The first considers linguistic diversity as a value in and of

itself that is to be preserved and promoted. This is, for instance, the main goal of the *European Charter for Regional or Minority Languages*. A second trend advocates for the strengthening of favorable measures for the dominant language of each State. Thus, in recent decades, more countries have adopted legislative measures to protect or ensure the internal and external presence of their majority language (France, Latvia, Slovakia, some states in the USA, and so on). Finally, there is also a tendency to gradually establish a *lingua franca* at the supranational level, although the preeminence of English has happened more as a result of the practical development of public (and private) policies (i.e. the European Union's academic programs) than of the adoption of formal legal regulations. All three of these trends are present to some degree in the current production of legal regulations.

The importance of language regulations does not derive only from the communicative function of languages; interactions between any given language and its social environment (or language ecology, as defined by Einar Haugen in 1972) are also highly relevant. Language is an essential component of human identity (Hogan-Brun and Wolff 2003: 3; Patten 2003: 313). Losing it, or facing the inequality and/or decline of one's own language results in personal and collective traumas, and is likely to cause social conflict (Kontra 1999: 281). Language loss or linguistic genocide (Skutnabb-Kangas 2000) are not merely linguistic issues, but are related to power and discrimination (May 2012: 4). Some of these power dynamics related to the presence and/or evolution of languages in society arise directly or indirectly from the legal regulations on this matter (or the lack thereof). Linguistic regulations are but the legal expression of the language policies decided on by each country.

Unlike religion, language remains a necessary element in exercising many of the functions of any State (Patten 2003: 296; Rubio-Marín 2003: 55). This results in the adoption of linguistic policies, overt and covert, in each State. The relationship between languages and the construction of national identity is the subject of numerous multidisciplinary studies (Judt and Lacorne 2004; Kamusella 2009). According to Spolsky (2003), the following factors determine the language policy of each State: national ideology, the globalization process, the sociolinguistic situation, and the interest in minorities. The legal contribution to analyses of language policies becomes essential, as the linguistic options for each State translate into legal rules, and these in turn have an impact on the linguistic dynamics of society.

From a legal perspective, the most critical political decision that can be made to promote a language is to declare it as an official State language. Making a language official involves a legitimization and institutionalization process (May 2012: 161), as well as a public recognition of the language as part of the collective identity (Kymlicka and Patten 2003: 6). This explains why the vast majority of European states have incorporated provisions regarding the use of their languages into their Constitutional Law (Ruiz Vieytez 2005). In this sense, it is necessary to study the relationship between the linguistic regime of a given State and this State's tradition, or in other words, its 'politics of language policy' (Cardinal and Sonntag 2015).

However, language contact is not limited exclusively to the public and state levels. The language contact areas that can be affected by specific regulations are much broader than that. On the one hand, language policies are also conditioned by multilevel governance, from the local to the international level. Contact between languages also requires regulations from international organizations, which in turn repeat the model by allowing

a small number of official (or working) languages. Among them, the European Union provides a complex response to language contact, as EU language regulations reflect the complexity of its political configuration (Arzoz 2008; Nic Shuibhne 2002; Van Els 2006; McAuliffe 2012).

Moreover, besides the public sphere, there are also other kinds of regulations related to language contact phenomena. In practice, all social and economic groups have a particular policy or means of 'language management' (Spolsky 2009), which may have an impact on internal rules and sanctions, and therefore, on legal rules. The most important phenomenon in this field nowadays is that of multinational companies, in which linguistic contact has become a focus of attention, both internally and regarding the markets in which they operate (Cameron 2011; Sherman, Engelhardt, and Nekvapil 2012).

4. Language contact, law, and justice

In spite of language contact being a widely spread phenomenon today, there is a prevailing social and political tendency to conceive of political spaces as monolingual (Kochenov and De Varennes 2014: 56). Most states in the world have only one official language, and a very high percentage of their inhabitants share that language. This in turn explains the fact that only an extremely small number of the world's languages are officially recognized (May 2012: 5; Romaine 2002: 1). The idea of monolingual nation-states, and the linguistic policies deriving from that idea, pose a serious legal threat to linguistic diversity (Kontra 1999: 284; Kymlicka and Patten 2003: 10).

Given the disparate political, legal, and symbolic weight of languages at the local, regional, state, and suprastate level, the world is full of situations of asymmetric multilingualism that must be managed. Population movements and globalization increase language contact, and many languages of the world are at risk of becoming extinct (Wurm 2001), while simultaneously, English has in fact been adopted and acknowledged as a worldwide *lingua franca*. The preeminence of some languages over others produces unfair effects of various kinds. This is where the concept of Linguistic Justice emerges (De Schutter 2007), aiming to secure the fair relationships that Law and languages should maintain in each specific context and in international society (Mowbray 2012: 144). In this Linguistic Justice debate (Alcalde 2015), the Law should play a prominent role, at least instrumentally.

The need to develop theories and policies based on the concept of Linguistic Justice is bolstered by a combination of several factors. The factors that are most often referred to include the consolidation of English as a global *lingua franca* (May 2012: 206–207), which generates differences in opportunities for speakers of this language when compared to speakers of other languages (Grin 2005: 455–456), and a sort of linguistic imperialism (Phillipson 2003). In order to properly compensate for this globalizing trend, various measures have been proposed, including 'language taxes' for English-speaking countries, a prohibition on dubbing audio-visual materials for general consumption, and strengthening of the territorial principle in linguistic regulations (Castiglione and Longman 2007: 226–247; Van Parijs 2011).

Yet linguistic injustice also hinges on the existence of nation-States and the idea that each State should have a common language. This ideology generates new linguistic

conflicts (Skutnabb-Kangas and Phillipson 1995) that can only be prevented through more inclusive and pluralistic language policies (May 2012). Existing International Law does not advance Linguistic Justice, since it does not pay enough attention to specific contexts, their complexity, and the systematic disadvantages facing many linguistic communities (Mowbray 2012: 202–205). In addition, International Law on minorities normally tends to depoliticize languages, in the sense that they are not viewed as identity-building elements, but rather as cultural products (Mowbray 2012: 150).

Linguistic Justice also has much to do with legal interpretations of the equality principle, and its implementation in settings characterized by language contact. Whereas some authors defend a per capita principle of linguistic justice (i.e. prorating language recognition based on a per capita distribution), others prefer the equal services principle (i.e. upholding official languages, irrespective of the number of speakers) (De Schutter 2016). Particular contexts and existing constitutional solutions may also influence positions in this debate, in which more voices of legal scholars would be desirable.

5. References

Alcalde, Javier
2015 Linguistic justice: An interdisciplinary overview of the literature. *Amsterdam Working Papers in Multilingualism* 3. 27–96.

Anderson, Benedict
1983 *Imagined communities: reflections on the origin and spread of nationalism*. London: Verso.

Arzoz, Xabier (ed.)
2008 *Respecting linguistic diversity in the European Union*. Amsterdam: John Benjamins Publishing Company.

Balogh, Arthur de
1930 *La protection internationale des minorités.* Paris: Editions Internationales.

Beukes, Anne-Marie
2009 Language policy incongruity and African languages in postapartheid South Africa. *Language Matters* 40(1). 35–55.

Cameron, Deborah
2011 Regulating language in the global service industry. In Catrin Norrby & John Hajek (eds.), *Uniformity and diversity in language policy*, 258–272. Bristol: Multilingual Matters.

Cardinal, Linda & Selma Sonntag (eds.)
2015 *State traditions and language regimes*. Montreal: McGill-Queens University Press.

Castiglione, Dario & Chris Longman (eds.)
2007 *The language question in Europe and diverse societies: Political, legal, and social perspectives*. Oxford-Portland: Hart Publishing.

De Schutter, Helder
2007 Language policy and political philosophy: On the emerging linguistic justice debate. *Language Problems and Language Planning* 31(1). 1–23.

De Schutter, Helder
2016 Two principles of equal language recognition. *Critical Review of International Social and Political Philosophy* 20(1). 75–87.

De Varennes, Fernand
1996 *Language, minorities, and human rights*. The Hague: Martinus Nijhoff.

De Varennes, Fernand
2001 Language rights as an integral part of human rights. *UNESCO International Journal on Multicultural Societies* 3(1). 15–25.
De Witte, Bruno
1985 Linguistic equality: A study in comparative constitutional law. *Llengua i Dret* 6. 43–126.
Extra, Guus, Massimiliano Spotti & Piet Van Avermaet (eds.)
2009 *Language testing, migration, and citizenship: Cross-national perspectives on integration regimes*. New York: Continuum International Publishing Group.
Fernández Liesa, Carlos
1999 *Derechos lingüísticos y derecho internacional.* Madrid: Dykinson.
Fouques Duparc, Jacques
1922 *La protection des minorités de race, de langue et de religion: Étude de droit des gens*. Paris: Dalloz.
Grin, Francois
2003 *Language policy evaluation and the European Charter for Regional or Minority Languages*. New York: Palgrave.
Grin, Francois
2005 Linguistic human rights as a source of policy guidelines: A critical assessment. *Journal of Sociolinguistics* 9(3). 448–460.
Guillorel, Herve & Genevieve Koubi (eds.)
1999 *Langues et droits: Langues du droit, droit des langues*. Brussels: Bruylant.
Haugen, Einar
1972 *The ecology of language: Language science and national development*. Stanford: Stanford University Press.
Henrard, Kristin
2003 Devising an adequate system of minority protection in the area of language rights. In Gabrielle Hogan-Brun & Stefan Wolff (eds.), *Minority languages in Europe: Frameworks, status, prospects*, 37–55. New York: Palgrave Macmillan.
Hogan-Brun, Gabrielle & Stefan Wolff
2003 Minority languages in Europe: An introduction to the current debate. In Gabrielle Hogan-Brun & Stefan Wolff (eds.), *Minority languages in Europe: Frameworks, status, prospects*, 3–15. New York: Palgrave Macmillan.
Howard, Martin
2007 *Language issues in Canada: Multidisciplinary perspectives*. Newcastle: Cambridge Scholars Publishing.
Judt, Tony & Denis Lacorne (eds.)
2004 *Language, nation and state.* New York: Palgrave Macmillan.
Kamusella, Tomasz
2009 *The politics of language and nationalism in modern Central Europe*. New York: Palgrave Macmillan.
Kochenov, Dimmitry & Fernand De Varennes
2014 Language and law. In Francis Hult & David Johnson (eds.), *Research methods in language policy and planning*, 56–66. Oxford: Wiley-Blackwell.
Kontra, Miklos
1999 Some reflections on the nature of language and its regulation. *International Journal on Minority and Group Rights* 6(3). 281–288.
Kymlicka, Will & Alan Patten
2003 Language rights and political theory: Context, issues, and approaches. In Will Kymlicka & Alan Patten (eds.), *Language rights and political theory*, 1–51. Oxford: Oxford University Press.

Mac Golla Chríost, Diarmait
2016 *The Welsh language commissioner in context: Roles, methods and relationships*. Cardiff: University of Wales Press.
May, Stephen
2012 *Language and minority rights: Ethnicity, nationalism and the politics of language*. New York: Routledge.
McAuliffe, Karen
2012 Language and law in the European Union: The multilingual jurisprudence of the ECJ. In Lawrence M. Solan & Peter M. Tiersma (eds.), *The Oxford handbook of language and law*, 200–208. Oxford: Oxford University Press.
Milian i Massana, Antoni
1994 *Derechos lingüísticos y derecho fundamental a la educación: Un estudio comparado: Italia, Bélgica, Suiza, Canadá y España.* Madrid: Civitas.
Mowbray, Jacqueline
2012 *Linguistic justice: International law and language policy.* Oxford: Oxford University Press.
Nic Shuibhne, Niamh
2002 *EC law and minority language policy: Culture, citizenship and fundamental rights.* The Hague: Kluwer Law International.
Nogueira, Alba, Eduardo Ruiz Vieytez & Iñigo Urrutia Libarona (eds.)
2012 *Shaping language rights: Commentary on the* European Charter for Regional or Minority Languages *in light of the Committee of Experts' evaluation.* Strasbourg: Council of Europe Publishing.
Patten, Alan
2003 What kind of bilingualism? In Will Kymlicka & Alan Patten (eds.), *Language rights and political theory*, 296–321. Oxford: Oxford University Press.
Phillipson, Robert
2003 *English-only Europe? Challenging language policy.* London: Routledge.
Pizzorusso, Alessandro
1993 *Minoranze e maggioranze*. Torino: Einaudi.
Reaume, Denise
2003 Beyond personality: The territorial and personal principles of language policy reconsidered. In Will Kymlicka & Alan Patten (eds.), *Language rights and political theory*, 271–295. Oxford: Oxford University Press.
Romaine, Suzanne
2002 The impact of language policy on endangered languages. *UNESCO International Journal on Multicultural Societies* 4(2). 194–212.
Rubio-Marín, Ruth
2003 Language rights: Exploring the competing rationales. In Will Kymlicka & Alan Patten (eds.), *Language rights and political theory*, 52–79. Oxford: Oxford University Press.
Ruiz Vieytez, Eduardo
2001 The protection of linguistic minorities: A historical approach. *UNESCO International Journal on Multicultural Societies* 3(1). 5–14.
Ruiz Vieytez, Eduardo
2005 Lenguas y constitución: Una visión del derecho lingüístico comparado en Europa. *Revista Vasca de Administración Pública* 72. 231–275.
Ruiz Vieytez, Eduardo
2009 Constitutions, languages, definitions, and the *European Charter for Regional or Minority Languages*. *Llengua i Dret* 51. 227–253.
Ryan, Kerry & Tim Mcnamara
2011 Testing identity: Language tests and Australian citizenship. In Catrin Norrby & John Hajek (eds.), *Uniformity and diversity in language policy*, 180–194. Bristol: Multilingual Matters.

Sherman, Tamah, Olive Engelhardt & Jiri Nekvapil
2012 Language use in multinational companies in Europe: A theoretical and methodological reframing. In Patrick Studer & Iwar Werlen (eds.), *Linguistic diversity in Europe: Current trends and discourses*, 287–310. Berlin: De Gruyter Mouton.
Skutnabb-Kangas, Tove
2000 *Linguistic genocide in education – or worldwide diversity and human rights?* Mahwah & London: Lawrence Erlbaum Associates.
Skutnabb-Kangas, Tove & Robert Philipson (eds.)
1995 *Linguistic human rights: Overcoming linguistic discrimination.* Berlin & New York: Mouton de Gruyter.
Spolsky, Bernard
2003 *Language policy.* Cambridge: Cambridge University Press.
Spolsky, Bernard
2009 *Language management*. Cambridge: Cambridge University Press.
Turi, Joseph
1990 Panorama comparé des différentes législations juridico-linguistiques en matière de politique linguistique: Les cas les plus intéressants. In Edorta Cobreros (ed.), *Jornadas sobre el régimen jurídico del Euskera*, 39–66. Oñati: Instituto Vasco de Administración Pública.
Turi, Joseph
1994 Typology of language legislation. In Tove Skutnabb-Kangas & Robert Philipson (eds.), *Linguistic human rights: Overcoming linguistic discrimination,* 111–120. Berlin & New York: Mouton de Gruyter.
Turi, Joseph
2015 Language law and language rights. *The International Journal of Law, Language & Discourse* 5(2). 1–18.
Van Els, Theo
2006 The European Union, its institutions and its languages: Some language political observations. In Richard B. Baldauf & Robert B. Kaplan (eds.), *Language planning and policy in Europe*, vol 2, 202–256. Clevedon: Multilingual Matters.
Van Oers, Ricky, Eva Ersboll & Dora Kostakopoulou (eds.)
2010 *A redefinition of belonging? Language and integration tests in Europe*. Leiden: Brill.
Van Parijs, Philippe
2011 *Linguistic justice for Europe and for the world*. Oxford: Oxford University Press.
Vernet, Jaume (ed.)
2003 *Dret lingüístic.* Valls: Cossetània.
Viaut, Alain
1999 Reconnaissance légale et représentations de la langue en situation de contact linguistique (autour des notions de loyauté et de prestige). In Herve Guillorel & Genevieve Koubi (eds.), *Langues et droits: Langues du droit, droit des langues*, 31–56. Brussels: Bruylant.
Willemyns, Roland
1992 Linguistic legislation and prestige shift. In Ulrich Ammon & Marlis Hellinger (eds.), *Status change of languages*, 3–16. Berlin: Walter de Gruyter.
Woehrling, Jean Marie
2005 *The* European Charter for Regional or Minority Languages*: A critical commentary.* Strasbourg: Council of Europe Publishing.
Wurm, Stephen (ed.)
2001 *Atlas of the world's languages in danger of disappearing*, 2nd edn. Paris: UNESCO Publishing.

Eduardo J. Ruiz Vieytez, Bilbao (Spain)

60. Contact linguistics and literary studies

1. Introduction
2. Philology's oddly modern dispreference for language contact
3. Picking up where they left off
4. Against linguistic tourism
5. Literature on the field of monolingualism
6. Translation as language contact
7. References

> Leopards break into the temple and drink to the dregs what is in the sacrificial pitchers; this is repeated over and over again; finally it can be calculated in advance, and it becomes a part of the ceremony. [*Leoparden brechen in den Tempel ein und saufen die Opferkrüge leer; das wiederholt sich immer wieder; schließlich kann man es vorausberechnen, und es wird ein Teil der Zeremonie*.] (Kafka 1969: 93)

1. Introduction

Any reasonable observer might expect literary studies and contact linguistics to strike up lively conversations with one another at every turn. After all, literary studies is at its best − isn't it? − when concerned with the rich, variable modes of language use cultivated in a range of aesthetic formats and creative settings, and thus also with how the social and sociolinguistic world can (and cannot) be brought to life on the pages of novels, short stories, and poems. Yet over the last two and a half centuries, literary studies, and in particular the national philologies, have crafted ways to avoid, or at least sequester, the question of language(s) in contact − and sometimes indeed the question of language(s) in general. Consider the bafflement of Corngold (2001: 281), writing admiringly about the (multilingual and contact-linguistically riveting) Kafka ([1919] 1995) story "In the Penal Colony": "[t]here is a good deal of difference between how this story has come down to us, as a systematically pure and grave conceptual meditation on the one hand; and, on the other, how it actually reads".

Like a musicologist who no longer quite hears the music, even the most passionate and critical literary analyst can get so wrapped up in themes, motifs, ideologies, and figures as to lose sight − for a moment, or even for one's whole career − of language, whether in the sense of *langage* (the unique human faculty of language and the idioms immanent in this faculty) or *langue* (any one, or a combination, of the thousands of particular semi-sovereign language systems used in this, or other potential, worlds). Perhaps it was this, among other occupational hazards of literary criticism, that made it necessary in the 1970s for such a reconciliatory book as Traugott and Pratt's (1980) *Linguistics for Students of Literature* to appear, the title of which rather gingerly brokers this proposed reunion between literature and linguistics. Indeed, around 1980, a nascent flurry of contributions fueled by linguistically informed literary stylistics and pragmatics, as well as of speech act theory and variationist sociolinguistics, made its presence felt in mainstream literary critical journals (notably Sternberg 1981).

https://doi.org/10.1515/9783110435351-060

2. Philology's oddly modern dispreference for language contact

In this way, the years around 1980 were an evanescent, but promising, moment in which multilingualism and linguistic contact were becoming a vigorously conceptualized topic in major venues like *Poetics Today.* Yet both of these insurgent scholarly developments − the multilingual and the contact/variationist linguistic − receded from literary discourse by 1990, as philosophical, cultural, and deconstructivist conceptions around Language (writ large) held sway. One result of this recession was that the underlying and unmarked credibility of (mono)lingualism persisted as a signature paradigm for the study of literature, whether national, transnational, or world-literary in its intension. Amid this conjuncture, Derrida's (1996) interventions in *Monolingualism of the Other,* about the unrelenting colonial myth of hexagonal French, contributed more to the legacy and pathos of Critical Theory in literary and cultural studies than it did to reanimating the dialogue Traugott and Pratt had sought to broker between linguistics and literary studies nearly two decades prior. Even as hybridity, interculturalism, postcolonialism, and globalization enjoyed more and more airtime in the literary disciplines over the 1990s, the problem of particular *langue(s)* (not to mention of linguistic nativism and protectionism; Chow 2014) remained muted amid these other urgent endeavors to 'rethink' literature for a global scale.

In the United States, neoconservative animus about bilingual education had tended since the 1980s to pre-politicize any question around language contact as it might pertain to literature or other creative aesthetics, and this aura around bilingualism's political topicality helped hamper its general reintroduction into the tool chest of literary aesthetics. Unabashed antimulticultural rancor (most notably expressed in Harold Bloom's [1995] *The Western Canon*) colluded with both monolingualist New Critical habits of literary analysis and a trade publishing industry increasingly disinclined to contract sociolinguistically variant literary projects whose intratextual multilingualism might complicate translatability for global market distribution. Phenomena of language contact, along with bilingualism and code-blending, were remanded yet again to the purview of sociology and educational linguistics, in part because the structuralist (and indeed even poststructuralist) legacies of literary criticism had presumed authorial monolingualism as their default construct, despite Derrida's later-stage intentions to divest from it. As the Romance comparatists Hokenson and Munson (2007: 148) put it, "[b]ilinguality seems to be the one category of language user that high modernist thought did not, indeed perhaps even refused to, consider".

3. Picking up where they left off

As noted, the early 1980s housed some adventurous retheorizations of the mono/multilingual paradigm in literary criticism, and leading voices in Chicanx feminist criticism (Gloria Anzaldúa, most prominently) gave expression to this moment of (re)discovery and (re)centering upon the 'tongue' in literary aesthetics. Since around 2012, a new literary movement around the notion of a 'philology of multilingualism' has also been working to reinvigorate the ways literary theory can countenance more than one language at once − and thus, how scholars and readers might more meaningfully unmoor from

their disciplines' allegiances to (multiculturally inflected) monolingual national canons. Several scholars, writing primarily in German and against the backdrop of *Germanistik*, have highlighted for critique the hereditary discomfiture arising from Ferdinand de Saussure's francophone conceptual spectrum of *langage-langue-parole* − which, these scholars point out, translates poorly into English and German, and had been only notionally elaborated in Saussure's most prominent writings to begin with. Stockhammer (2017) observes furthermore how the 'linguistic turn' and 'language games' associated with Rorty and Wittgenstein, respectively, are grounded in a conceptualization of *langage* rather than *langue* − i.e. that they are primarily concerned with the faculty and condition of Language in the general humanistic sense, rather than with language in its plural, particular forms.

Rath (2017) makes a similar argument about the famed *fin de siècle* Central European 'crisis of language', as expressed in Hugo von Hofmannsthal's canonical 1902 literary experiment "A Letter" (which itself long went under-recognized as a multilingual pseudo-translation from English to German). Rath shows how the linguistic-crisis discourse, touted forebodingly in the literary criticism around this key text, has long been predicated both on a monolingualist misapprehension of "how it actually reads" (Corngold 2001: 281) and a misplaced longing for the idealized faculty of Language as such, rather than on a sober and worldly encounter of the practical (multi)linguality as it is lived and spoken. This enduring equivocation around *langage* and *langue* in literary critique, Stockhammer (2017) claims, is not often enough recognized as a paradigm problem, stemming as it does from the all-too-apropos translational predicament that French can speak easily and well of a difference between *langage* and *langue*, while English and German cannot − try as they might. Hofmannsthal and Wittgenstein (with their '*Sprache*') and Rorty (with his 'language') cannot quite access, in a reproducible and systematic way, the conceptual range that French offers.

Among the remedies literary theorists of this emerging 'Luxemburg School' have forged to highlight the *langage-langue* problem is a distinction between so-called *Sprachlichkeit* and *Sprachigkeit* ('linguality' and 'lingualism', respectively), a binary that critically plies at the widely circulating and largely unquestioned term *Mehrsprachigkeit* ('multilingualism', see also Martyn 2014). This peculiarly modern and Global-Northern sense of *Mehrsprachigkeit* − as a multiplication of lingual*ism*s or monolingualisms (see Pennycook 2010) − belies its *-ig* suffix in German and, in turn, its *-ism* suffix in English, to the extent that multilingualism in contemporary political, cultural, and social discussions is presumptively understood as an additive accomplishment, resulting from the acquisition and accumulation of complete, fortified language integers like English, German, French, Tagalog, Egwe, Diné, etc. In order to enter into the symbolic order of contemporary 'multilingualism', a speech repertoire must be or become countable as a *langue* (see Duchêne and Heller 2012; Makoni and Pennycook 2006), and speakers must also possess and demonstrate such integers in order to count as multilingual in this order. The late eminent Bakhtin translator Holquist (2014: 8) put a fine point on this when he claimed that certain predominant forms of multilingualism have been constituted on the back of what he called 'linguistic monism', or: "a positive belief system [that] conceives the world as consisting of geographically dispersed common languages each of which has a unique separate identity of its own that is both stable and unitary".

Holquist (2014) indeed advanced this critique of 'linguistic monism' in the late years of his career, after decades of efforts to expose anglophone readers to Mikhail Bakhtin's

vision of the types of heteroglossia characteristic of the genre of the novel. In his essay "What Would Bakhtin Do?", Holquist further sought to demonstrate that such a critique of 'lingualism' had been at the heart of Saussure's endeavor all along, especially in his late work, where Saussure develops the following metaphor (as cited in Holquist 2014: 13):

> A sign system must be part of a community. Indeed, any semiological system is not a ship in dry dock, but a ship in the open sea. Which is the real ship: one in the covered yard surrounded by engineers, or a ship at sea? Quite clearly only a ship at sea may yield information about the nature of a ship. A community environment changes everything. A sign system is destined for a community just as a ship is destined for the sea.

To connect Saussure's metaphor to the currently emerging critical lexica of 'multilingual philology', we may speak of linguality (*Sprachlichkeit*) as the 'ship at sea', i.e. the ship in its relation to its community of use, while lingualism is that ship in 'dry dock', analytically isolated from the conditions and wear of use. In a similar vein, Steffensen and other ecological thinkers about language offer the helpful distinction between a so-called 4E model of 'distributed language use' (i.e. language as always embodied, embedded, enacted, and ecologically oriented), as opposed to 3I models of language (as internal, instrumental, and individual; Steffensen 2015). Naturally, the composite model of 4E linguality 'at sea' is much more methodologically appropriate for countenancing complex and constant language contact phenomena in literature and elsewhere, than is that of 3I 'dry dock' lingualism. And yet, the latter model still holds sway over most methods of literary criticism, not to mention policy platforms and language learning curricula.

For Holquist, inspired by Saussure and Bakhtin, literature was nevertheless a potentially instructive − if not revelatory and crucial − *locus amoenus* for such a community of multilingual living language use 'at sea'. Tempting as it is to then expect that there would inevitably be a fast friendship between literary studies and contact linguistics, such an expectation ignores some of the consistent historical facts about the practice and industry of literature since the 17th century. Scandalous and counterintuitive as it might sound, modern Global North 'literature' has primarily been the proving ground of *mono*lingualism, rather than *multi*lingualism. As Anderson ([1983] 2006) theorized now 35 years ago, the 19th century novel was not only a canvas for artistic expression of complex livelihoods in motion, but was also a new and effective technology for instantiating the nation as a community of mutual comprehension, 'homogenous, empty space', and monolingual contiguity. Where and when the (European) novel thrived, linguistic contact was likely to suffer direct, negative consequences and elisions. The advent of the novel as the binding aesthetic artifact of the reading 'imagined community' of the nation has augured ill for social multilingualism and language contact over the course of the golden age of nation-building.

Commercially sanctioned forms of literary production over this period became some of the most charismatic 'Trojan horses' for ideological monolingualism and linguistic nationalism, cloaked as they were in the garb of aesthetic exploration and cosmopolitan progress. Consolidating these forms in the 20th century, trade publishing houses made the monolingual, translatable book their primary and prohibitive mode of commercial salience (Komska 2018; Lennon 2011). As literary markets began to extensively global-

ize in the 1980s, the market imperative to acquire and option quintessentially translatable books for a roster of imagined end-user monolingual markets, in accordance with the model of 'lingualism' described above, began in turn to restructure the production of literature across borders (Rectanus 1987). Literature, which had served symbolically as the bellwether of cultural particularity in the 19th and early 20th centuries (e.g. of 'the American experience' or of 'the Russian soul') was transformed in the 1990s into a new mandate: one must overcome particularity so as to convey content across linguistic borders (Parks 2010; see also Noorani [2013] on earlier forms of 'soft multilingualism' in Arab nationalist poetry).

4. Against linguistic tourism

Still, in the face of nationalistic (or otherwise ideological) monolingualism, literary texts have long been able to serve, at least potentially, as a safe harbor for the 'translingual *imagination*' (Kellman 2000, emphasis added), hosting not only the carnivalesque heteroglossia of living languages, but also offering fertile territory for the production of new forms of language − forms that may still be unthinkable in other social and institutional realms. Literary creators from Yoko Tawada to Junot Díaz have become known as master code-mixers, bringing readers into the translingual 'traffic in meaning' (Pratt 2002) and engaging them in the 'arts of the contact zone' (Pratt 1991). A host of literary studies-based endeavors, ranging from Wiggin and McLeod's (2016) collection *Un/Translatables* and the mammoth undertakings of Cassin (2004) in her collaboratively produced *Vocabulaire européen des philosophies: Dictionnaire des intraduisibles*, to Sommer's (2003) *Bilingual Games,* have taken up the challenge of translingual aesthetics and imagination as an object of inquiry and critique. Moreover, as the definition of 'literature' itself evolves (bearing in mind how the category 'literature', in the sense European languages know it, has only existed only for a few centuries), 'literariness' is also today growing to embrace oral, improvisational, multimodal, and hybrid genres of creative practice.

Such paradigm shifts have captured the attention of some academic traditions since 1990, a trend that has led literary and cultural studies over its most recent generations to highlight, and occasionally to overstate, the role of language contact in literary cultures whose roots lie in European elite models. Matsuda (2014: 478−483), for instance, cautioned against the "lure of translingual writing", claiming that the translingual has become a kind of darling of literary thinking over recent years, and that this attraction to multilingual literary figuration both undermines and paternalizes the intellectual projects at work therein. Applied linguists like Matsuda seem to be suggesting that literary scholarship first needs to forthrightly address its own historical, formal, and methodological indebtedness to monolingualism, before celebrating manifest code-switching and language contact as these may be occasionally foregrounded in published literary work.

5. Literature on the field of monolingualism

Certainly, earlier generations of literary creators also harbored an acute sense of the attenuating forces of monolingualism (in its ideological and structural forms) on their

craft. Kafka (1990: 818), for instance, wrote in a diary entry in 1917 that: "[l]iterature, to express it in reproachful terms, is such a drastic curtailment of language. […] The noise-trumpets of nothingness [*Literatur, als Vorwurf ausgesprochen, ist eine so starke Sprachverkürzung.* […] *Die Lärmtrompeten des Nichts*]". Generations of Kafka scholarship have assumed that his exasperated anguish about language arose out of his experiences of the finitude of the human faculty of Language (i.e. Saussure's *langage*), and due to the vexingly meager transposability of the world of thought and symbolic experience onto the pages of printed, published text. In the last 25 years, as scholarship on Kafka has more and more come to rediscover his multilingualism in Czech, Yiddish, Hebrew, French, and English, scholars have begun to seriously consider how his command of these additional languages caused Kafka quite practical problems in a context where his literary production was compelled to acquiesce to the normative monolingualism of his age, its generic literary legacies, and the sociolinguistically narrow conceptions of comprehensibility harbored by his publishers (Čermák 1994).

The literary critic Wirth-Nesher (2008) provides another early example of the ways translingual authors found it necessary, and even productive, to adapt to a literary-industrial world that promotes (mono)lingualism over (multi)linguality. Mary Antin (1881–1949), one of the English-writing exophone authors Wirth-Nesher surveys in her monograph *Call it English,* puts it this way: "I was born, I have lived, and I have been made over. Is it not time to write my life's story? I am as much out of the way as if I were dead, for I am absolutely other than the person whose story I have to tell" (as cited in Wirth-Nesher 2008: 50). There is a certain sovereign haughtiness in this stance, as Antin further projects: "I shall turn the aliens' ridicule into sympathy. This I can do, for I am both of you and of them. I speak both your languages" (Wirth-Nesher 2008: 52). Writing (in English) becomes a sacred ceremonial goal, one that requires her to align herself with the requirements of a tradition that is not her own.

The lingualism/linguality divide in literary production is a particularly troubling dynamic for writers who grew up with an adverse relationship to the 'lingualism' presumed of them, for instance, those who, as Kerrigan (2008: 10) writes in *Archipelagic English*, "are anglophone but not necessarily English and who often have no place in the canon, or, more insidiously, are there on the wrong footing". Kerrigan thus calls for a 'devolution' of English in order to rehistoricize Hiberno-English and Scots, a prospect similar in scale and zeal to Makoni and Pennycook's (2006) call for 'disinventing languages'. As Kerrigan (2008: 2) writes: "[t]he incentive to strip away modern Anglocentric and Victorian imperial paradigms to recover the long, braided histories played out across the British-Irish archipelago between three kingdoms, four countries, divided regions, variable ethnicities and religiously determined allegiances is there even for those who are skeptical about the desirability of Scottish or Welsh independence".

While that project of a linguistically devolved philology of multilingualism is underway, authors working from Scottish or Hiberno-English will be adversely positioned in the diplomatic system of monolingualisms that has long counted their texts as part of English literature. For Jackson (2009: 4–5), this means that "the novel is the most alphabetic kind of written story", in that it requires a constant assimilatory tack toward the metropole's lingualism. The Scottish novelist James Kelman describes what happens when, for instance, Scots semiodiversity is sought in a novel, beyond what Jackson considers the metropole's *alphabetography.* As Kelman (2008: 60–61) writes:

> [i]n Scottish the 'man' will speak with what is called a 'heavy burr'. This is what the general run of lower-order Scottish people have in English literature, 'heavy burrs'. The writers may highlight this within the text. Colonial servants and underlings are integral to English literature and different conventions exist to deal with them. One such convention is the apparent attempt at phonetic transcription. I mean by that the spelling of words to give the impression of sound. I say 'apparent' because there is no authentic attempt going on.

Such are some of the literary problems that have emerged most intensively in a period we may choose to call multicultural, postethnic, superdiverse, or postmonolingual − a period since 1990 where ethnonationality has ceased to offer sufficient explanatory power about the nature or import of a literary work circulating beyond its immediate linguistic context. And yet, in most accounts of the system or endeavor of World Literature, a consideration of monolingualism and multilingualism plays only the most peripheral role. The language of composition is regarded as one participatory element in the cultural repertoire of literary creation writ large, and on that specific contextual basis is regarded as a challenge for translation or translatability. Yet the language of literary composition is seldom subjected to much historical analysis, as a medium which has necessarily been created, tempered, reduced, and aligned throughout a modern process of normative standardization in order to be enfranchised to appear in a given literary text at all. That a text is 'written in Dutch' reveals little about how the Dutch language became (and becomes) a unified language, able and entitled to do commerce with other peer monolingualisms on the world-literary, political, and technological field.

And yet literary texts often strain against their own compelled monolingualisms in unruly ways. Let us take the further case of a recent German-language novel called *Die Ohrfeige*, written by Abbas Khider (2016), a German citizen of Iraqi descent, a refugee from Saddam Hussein's jails and torture chambers, and a transmigrant who never intended to land in Germany permanently. *Die Ohrfeige* (or, *The Slap*) is a tale of immigration, asylum petition, noncomprehension, bureaucratic irrationality, and political limbo. Akin to predecessor texts like Kafka's (2008 [1911−1914]) *Amerika* novel, *The Slap* is written in German, but instantiates a storyworld that is entirely Arabic and understands little German. The narrator shares thoughts (in German), like: "[i]f I knew German, this is what I would say to you [ ]" (Khider 2016). The manifest text is therefore a testament to the multilingual experience, and to the minute, embodied endeavor of code-switching and linguistic contact, but only in relief: the artifact that one purchases and reads is entirely monolingual German, save a few cameo appearances of Arabic words, transliterated into Roman letters. Linguistic contact is thus indexed in this literary text as a political, epistemological, and critical double-bind *through its absence*, but the text does not itself *perform* linguistic contact in the sense of presenting variously and mutually incomprehensible languages in the same setting and format. This custom epitomizes mainstream literary production's engagement − indeed, its eager and desirous engagement − with linguistic contact 'themes', while procedurally and structurally diminishing the extent to which linguistic contact is able, in such settings, to emerge and enact moments of meaning-making.

Yet the cautionary example of *The Slap* does not end here. Audiobooks, often presumed to be a cloying and diminutive imitation of literary craft, step into the breach of multilingual and contact-linguistic phenomenality. The audiobook version of *The Slap*, produced in collaboration with the author, is − in Gerard Genette's ([1982] 1997) sense − a hypotext to the published literary text, in the sense of its derivative, displaced, compro-

mised, and nonautonomous literary standing. And yet, the audiobook is the place where (mono)lingualism emerges as truly an object of critique and creative disruption. While customary standard broadcast German is spoken by one of the audiobook readers, other reading voices in the audiobook − which at times sounds more like a radioplay − delight and dilate in multiethnolectal German pronunciations, and extended, overlapping utterances in Arabic. The multilayered nature of the audio tracks produces linguistic contact syntactically and paradigmatically, thus instantiating the kind of heteroglossic construction that Bakhtin theorized as the very nature of 'discourse in the novel'. And yet, such multilayeredness is preemptively excluded from the printed novel itself, while embraced and magnified in the audio version. The performance aesthetics in the audiobook is elevated and enthused amid the opportunity its reader-performers are seizing to imbue a patently multilingual story with actual performances of multilingualism and linguistic contact − that is, to return to the textworld the linguality that was always already properly its own.

6. Translation as language contact

While multilingual philology continues to forge new lines of thought around comprehension and incomprehension, one realm in which language contact in literature is fundamental, constant, and yet often rendered invisible, is literary translation (see also Chapter 64). The Translation Studies scholar Venuti (2016) has recently issued a challenge to literary comparatism to cast off its preferential treatment of 'original' texts over translations, and to cease 'suppressing' translated texts in their potential role as moments of linguistic contact, both in the classroom and in literary research.

The notion of translation as itself a contact-linguistic interaction has led to some inspired book-length studies. Kong's (2004) *Consuming Literature: Best Sellers and the Commercialization of Literary Production in Contemporary China*, demonstrates the various ways in which Western classics, often stylistically prohibitive in their own readerly contexts, are hyperdomesticated in Chinese popular literature markets, such that *Finnegans Wake*, for instance, is pitched ethnographically as an 'insightful look into the Irish mind'. In order to produce this ethnographic gaze on Western classics, Chinese translators disambiguate and simplify the English original, such that, for example, Molly Bloom's stream-of-consciousness narration is broken up into easily consumable phrases in Chinese, while remaining opaque and disobliging in the 'English' original, an English that through Joyce's compositional craft becomes a vehicle for anticolonial, anti-imperial, and even anti-anglophone critique. In such globalized translation processes, we see language-contact phenomena emerging in the form of an international literary idiom (regardless of the surface-level language of publication), in which the immanent critical resources of the language of composition are subsumed.

Some literary authors forthrightly take up the mediating contact-linguistic arena of literary translation as itself a critical element of their fictional projects. Kelman (2002), in his *Translated Accounts,* dramatizes the banality of collecting, recording, disseminating, and presenting genocidal brutality in a 21st century where human rights are tantamount to the right to have one's human rights translated and translatable. Kelman (2002: ix) introduces his novel with the following dispassionate introduction:

> These 'translated accounts' are by three, four or more individuals domiciled in an occupied territory or land where a form of martial law appears in operation. Narrations of incidents and events are included; also reports, letter-fragments, states-of-mind and abstracts of interviews, some confessional. While all are 'first hand' they have been transcribed and or translated into English, not always by persons native to the tongue. In a very few cases translations have been modified by someone of a more senior office. The work was carried out prior to posting into the computing systems. If editorial control has been exercised evidence suggests inefficiency rather than design, whether willful or otherwise. […] A disciplined arrangement of the accounts has been undertaken. Some arrived with titles already in place, others had none and were so assigned. Chronology is important but not to an overriding extent, variable ordering motions are integral to the process of mediation that occurs within computing systems and other factors were taken into consideration. It is confirmed that these accounts are by three, four or more anonymous individuals of a people whose identity is not available.

So begins the novel − with stacks and stacks of passive voice, and boilerplate, automated prose that records, but does not communicate, language contact and the tumultuous, violent transformations that often attend it. In a way, Kelman's prose here encapsulates many of the phenomena of linguistic contact that so thoroughly suffuse all literary worlds. Linguistic contact is presented and absented at the same time. It is sometimes described clinically, sometimes obscurely, sometimes stoically, ever desirously − and this mix of writerly affect often indexes the fundamental and profound predicaments that face any literary author in a multilingual world where what one means by 'a language' makes all the difference.

7. References

Anderson, Benedict
2006 [1983] *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.

Bakhtin, Mikhail
1981 *The dialogic imagination* [ed. Caryl Emerson; transls. Caryl Emerson & Michael Holquist]. Austin: University of Texas Press.

Bloom, Harold
1995 *The western canon.* New York: Riverhead Books.

Cassin, Barbara (ed.)
2004 *Vocabulaire européen des philosophies: Dictionnaire des intraduisibles.* Paris: Le Seuil & Le Robert.

Čermák, Josef
1994 Franz Kafkas Sorgen mit der tschechischen Sprache. In Kurt Krolop & Hans D. Zimmerman (eds.), *Kafka und Prag*, 59−66. Berlin: de Gruyter.

Chow, Rey
2014 *Not like a native speaker: On languaging as a postcolonial experience*. New York: Columbia University Press.

Corngold, Stanley
2001 Allotria and excreta in 'In the Penal Colony': For Rachel Magshamhrain. *Modernism/Modernity* 8(2). 281−293.

Derrida, Jacques
1998 [1996] *Monolingualism of the Other, or, The Prosthesis of Origin* [transl. Patrick Mensah]. Stanford: Stanford University Press.

Duchêne, Alexandre & Monica Heller (eds.)
2012 *Language in late capitalism: Pride and profit.* London: Routledge.
Genette, Gérard
1997 [1982] *Palimpsests: Reading the second degree* [transls. Channa Newman & Claude Doubinsky]. Lincoln: University of Nebraska Press.
Hokenson, Jan Walsh & Marcella Munson
2007 *The bilingual text: History and theory of literary self-translation*. Manchester: St. Jerome Press.
Holquist, Michael
2014 What would Bakhtin do? *Critical Multilingualism Studies* 2(1). 6–19.
Jackson, Tony E.
2009 *The technology of the novel: Writing and narrative in British fiction*. Baltimore: Johns Hopkins University Press.
Franz Kafka
1969 *Parables and paradoxes* [transl. Clement Greenberg]. New York: Schocken Books.
Kafka, Franz
1983 *Der Verschollene*. Frankfurt am Main: Fischer.
Kafka, Franz
1990 *Kritische Ausgabe: Tagebücher* [eds. Hans-Gerd Koch, Michael Müller & Malcolm Paisley]. Frankfurt am Main: Fischer.
Kafka, Franz
1994 *Kritische Ausgabe: Drucke zu Lebzeiten* [eds. Wolf Kittler, Hans-Gerd Koch & Gerhard Neumann]. Frankfurt am Main: Fischer.
Kafka, Franz
1995 *The complete short stories* [transls. Willa Muir & Edwin Muir]. New York: Schocken Books.
Kafka, Franz
1998 *The castle: A new translation, based on the restored text* [transl. Mark Harman]. New York: Schocken Books.
Kafka, Franz
2008 *Amerika: The missing person* [transl. Mark Harman]. New York: Shocken Books.
Kellman, Steven
2000 *The translingual imagination*. Lincoln: University of Nebraska Press.
Kelman, James
2002 *Translated accounts*. New York: Anchor Books.
Kelman, James
2008 '*And the judges said ...': Essays.* Edinburgh: Polygon.
Kerrigan, John
2008 *Archipelagic English: Literature, history, and politics 1603–1707.* Oxford: Oxford University Press.
Khider, Abbas
2016 *Die Ohrfeige.* Munich: Carl Hanser.
Komska, Yuliya
2018 Why Curious George did not speak: The conspicuous multilingualism of Margret and H. A. Rey. *German Studies Review* 41(3). 505–528.
Kong, Shuyu
2004 *Consuming literature: Best sellers and the commercialization of literary production in contemporary China*. Stanford: Stanford University Press.
Lennon, Brian
2011 *In Babel's shadow: Multilingual literatures, monolingual states*. Minneapolis: University of Minnesota Press.

Makoni, Sinfree & Alastair Pennycook
2006 *Disinventing and reconstituting languages*. Bristol: Multilingual Matters.
Martyn, David
2014 Es gab keine Mehrsprachigkeit, bevor es nicht Einsprachigkeit gab: Ansätze zu einer Archäologie der Sprachigkeit (Herder, Luther, Tawada). In Till Dembeck & Georg Mein (eds.), *Philologie und Mehrsprachigkeit*, 39–51. Heidelberg: Winter Verlag.
Matsuda, Paul Kei
2014 The lure of translingual writing. *PMLA* 129(3). 478–483.
Noorani, Yaseen
2013 Hard and soft multilingualism. *Critical Multilingualism Studies* 1(2). 7–28.
Parks, Tim
2010, 9 February The new dull global novel. *The New York Review of Books*. http://www.nybooks.com/daily/2010/02/09/the-dull-new-global-novel/. (accessed on 22 January, 2019).
Pennycook, Alastair
2010 *Language as a local practice*. London: Routledge.
Pratt, Mary Louise
1991 Arts of the contact zone. *Profession* 1991. 33–40.
Pratt, Mary Louise
2002 The traffic in meaning: Translation, contagion, infiltration. *Profession* 2002. 25–36.
Rath, Brigitte
2017 Speaking in tongues of a language crisis: Hugo von Hofmannsthal's *Ein Brief* as a non-monolingual text. *Critical Multilingualism Studies* 5(3). 75–106.
Rectanus, Mark W.
1987 Literary publishing in the Federal Republic of Germany: Redefining the enterprise. *German Studies Review* 10(1). 95–123.
Sommer, Doris
2003 *Bilingual games: Some literary investigations*. New York: Palgrave Macmillan.
Sommer, Doris
2004 *Bilingual aesthetics: A new sentimental education*. Durham: Duke University Press.
Steffensen, Sune Vork
2015 Distributed language and dialogism: Notes on non-locality, sense-making, and interactivity. *Language Sciences* 50. 105–119.
Sternberg, Meir
1981 Polylingualism as reality and translation as mimesis. *Poetics Today* 2(4). 221–39.
Stockhammer, Robert
2017 Converting lingualism into linguality (langagification des langues) in Goethe's Wilhelm Meister novels. *Critical Multilingualism Studies* 5(3). 32–51.
Traugott, Elizabeth Closs & Mary Louise Pratt
1980 *Linguistics for students of literature.* New York: Harcourt College Publishing.
Venuti, Lawrence
2016 Hijacking translation: How comp lit continues to suppress translated texts. *Boundary 2* 43(2). 194–204.
Wiggin, Bethany & Catriona MacLeod (eds.)
2016 *Un/Translatables: New maps for Germanic literatures.* Evanston: Northwestern University Press.
Wirth-Nesher, Hana
2008 *Call it English: The languages of Jewish American literature.* Princeton: Princeton University Press.

David Gramling, Tucson, Arizona (USA)

61. Media/Communication studies

1. Introduction
2. Development of media research from a language perspective
3. Researching the effect of the media on minority languages
4. Media as a tool for language maintenance
5. The impact of minority language media on the status and recognition of the language
6. A new ecosystem and the trilingual trap
7. Institutional completeness: a balancing act
8. Discussion
9. References

1. Introduction

Grounded in media studies, this chapter discusses how different minority languages are served by the media in various contexts. It specifically addresses how the needs of minority language speakers are fulfilled by the media in different languages. A pertinent issue in this context is how users of minority languages consume media in various languages in the emerging digital media ecosystem. Our focus on minority languages can contribute to many aspects of research on language contact, such as issues of language conflict, methodologies useful for studying language contact, language standardization, language shift, and language planning, policy, and politics. This focus is particularly interesting, as minority language speakers often have more varied language skills and needs (which they fulfill through their daily media habits) than the 'average' mainstream media consumer. Furthermore, minority languages tend to be more threatened than nurtured by the media, as a result of varied and asymmetric language contacts.

2. Development of media research from a language perspective

Current media research often relies on methods rooted in linguistics. The examples are manifold: close reading, discourse analysis, rhetoric, argumentation analysis, and others, all of which allow the researcher to analyze the content aspects of media. Studies combining media and language contact also frequently lean on policy research, including aspects of the regulatory frameworks that affect media supply and consumption and/or of the economic considerations that affect how media can reach different socioeconomic and linguistic environments (for a general overview, see Browne and Uribe-Jongbloed 2013). Interestingly, all these aspects are in a state of flux today, as the media are undergoing a profound process of change. This change could be called a second Gutenberg revolution: the evolution of a global sociodigital space that fosters new intersections between the human mind and virtual intelligence.

When placing language contact in the field of media research, it is useful to remember the role that the media have played in creating, promoting, and excluding linguistic communities. In Europe in particular, broadcasters assumed the role of national − indeed if not, of nation-building − institutions. Broadcasters became gatekeepers for languages spoken within these nations. This could include a recognition of multilingual realities (Moring and Husband 2007). Smaller language groups, however, such as the Sámi in

https://doi.org/10.1515/9783110435351-061

the Nordic states, were not automatically included in this national arena, but slowly had to work their way in.

In the 1970s and 1980s, terms such as 'cultural imperialism' (Demont-Heinrich 2011; Tomlinson [1991] 2001), 'dependency theory' (Ball-Rokeach and DeFleur 1976), 'marginalization', and 'new ethnicity' (Hall 1997) emerged. Western and globalized media were seen as spreading Western and US values and languages with little or no regard for local needs, thereby posing a severe threat to local cultures and languages. In the 1990s, however, the local and regional opportunities offered by the media began receiving more attention. For example, by introducing the concept of 'sphericules', Gitlin (1998) explained how regional cultures (such as Latin American or Egyptian) could create public spheres of their own based on a regionally shared language (Spanish or Arabic, respectively). The media could thus become an avenue for strengthening local and regional languages and cultures, instead of threatening them. This idea has been further strengthened by researchers of minority language and minority ethnic media through their perspectives on the role of the media in smaller linguistic communities (e.g. Cottle 2000).

In the current decade, however, we increasingly seem to be in a situation where language contact has become, if not a threat, then at least a double-edged sword. Recent studies have shown, for example, how controlling the 'proper' use of a national language has become difficult in the age of the Internet and social media (e.g. Blackwood 2013), and how English has become a normalized part of promoting media content in non-English-speaking communities (e.g. Hilgendorf 2013). Given that every individual can be a producer of media content today, the question becomes whether we are moving towards a reality of linguistic 'super diversity' within the media field.

However, it is important to emphasize that, even amongst minority language communities, realities differ dramatically. While some minority languages have access to an orthography recognized by technology as well as technological means, other languages may have neither. According to UNESCO, about one third of all languages of the world are endangered, and about half of these are severely or critically endangered (Moseley 2010). It thus cannot be assumed that media as a tool are available in equal measure to all minority language groups.

3. Researching the effect of the media on minority languages

The topic of whether, or how, media affect languages has been widely debated (Androutsopoulos 2014; Coupland 2014; Sayers 2014). Reports suggest that the media-language relationship is contingent upon broader societal relationships which need to be addressed in all their complexity. Explanations cannot be found in media alone (Coupland 2014; Sayers 2014). The operational conception of 'media' must be opened towards "an inclusive approach that encompasses institutions, technologies and practices of mediation as elements of the sociolinguistic condition of late-modern societies" (Androutsopoulos 2014: 247). Introducing the term "media linguistics", Cotter, Perrin, and Whitehouse (2017: 3) argue that "language and media research […] is necessary and essential to understanding language in its role in society and in terms of its creative − and limiting − potential". Among the many different functions that media/language dynamics can per-

form, Androutsopolous (2017) points to the particular ideological influence of the media on motivation(s) for minority language use, and to theoretical models and frameworks for understanding said influence.

The powerful effect of passive access to a second language through viewing television and/or playing video games has already been confirmed by many studies (Kuppens 2010). However, with browsing, information seeking, interacting through social media, and other forms of interpersonal contact-seeking on the Internet, the daily use of second language(s) has changed considerably. Media interfere in more and more dimensions of daily life, and especially in the lives of young adults. Time spent on various media is increasing rapidly. For instance, through multitasking on different devices, the total use-time of media and communications in the UK in 2014 had reached an average of over 11 hours daily, with even more alarming figures for the most media savvy cohort of 16–24-year-olds (Ofcom 2014).

Early authorities in the field of minority language retention, such as Fishman (2001: 473) and others, exercised considerable influence on interest in researching minority language media, having found that mass media tend to undermine endangered languages rather than support them. Fishman also aired some skepticism regarding the media's potential impact on language revitalization. Cormack (2007: 62) proposed a different route of inquiry that was based on media's negative impact, questioning "[i]n what ways [different media can] interact with other aspects of language use to contribute, directly or indirectly, to language maintenance in specific communities".

In response to an initiative of the Norwegian Sámi radio and television (*NRK Sápmi*), a desk study (unpublished) was carried out in 2014 by Tom Moring with Elisabeth Wide at the Swedish School of Social Science, University of Helsinki, which focused on different aspects of media effects on linguistic minorities. The study identified 75 articles that addressed the effects of media on language; however, hardly any were related to language contact. The articles discussed socialization and identity effects, as well as effects on second language learning, economic vitality, democracy, power, and language rights. They also discussed globalization, mediation, and mediatization, and the overall need for minority languages to develop digital media. All these themes were connected by a common thread: the fact that media are part of everyday language *use*, and this evidently has effects, however, said effects have not been identified in detail.

Many studies have suggested effects – either direct or indirect – of language contact on language *maintenance* and *construction*. For example, Swedish journalists in Finland have attempted to positively influence the minority language through broadcast journalism. They strongly subscribe to a 'standard language ideology', while at the same time opening up the concept of standard language to a broader dialectal diversity (Stenberg-Siren 2018: 3). Lehtola (1997: 12, 92) argues that Sámi media have guarded the Sámi language against the dominance of the majority language by elaborating the Sámi language, by giving it presence in the public domain, and by deliberately reintroducing vocabulary and expressions that have been close to extinction. Pietikäinen (2008) found that the Sámi media are part of a hybridization process which is characteristic of the media landscape today. Pietikäinen (2008) and Markelin, Husband, and Moring (2013) report similar views on potential language impact and ambitions held by journalists who work for Indigenous peoples' media. This issue is, as all of the aforementioned researchers note, highly politicized, in the sense of both the political economy surrounding the media environment, as well as the rights and resources afforded to Indigenous peoples.

4. Media as a tool for language maintenance

Media workers in language communities that contain many languages − such as the Sámi communities, with nine Sámi languages and four majority languages − find themselves in a situation where linguistic choices must be made every day. Markelin, Husband, and Moring (2013) reported that many members within the Sámi media see themselves as linguistic workers as much as media workers. Some senior editors' comments on their roles *vis-à-vis* the Sámi language are given below:

> Well for me it's the Alpha and the Omega. Because I see myself as an important person in the Sámi society − or not person, but role, I have an important role. And when I watch my children grow up: they have Sámi as their home language, they speak Sámi as their first language in school, and they are pretty good at using the little Sámi content we have in the vast digital world. But at the same time I see that their language is so impoverished. So to me, the Sámi media and NRK Sápmi in particular, is a very, very, important foundation stone in preserving the language and culture. (Markelin, Husband, and Moring 2013: 109–110)

> For me the Sámi language is the foundation for Sámi identity. I am one of those who is quite sure that if the language disappears then the Sámi identity disappears as well. […] And this is the driving force [for me]; I want Sápmi to exist in 50 years and in 100 years and so on into the future. And then it is important that the Sámi language is alive. (Markelin, Husband, and Moring 2013: 109)

Media are thus seen as an essential tool for language maintenance and revitalization. However, about 70 % of the Sámi people have lost their Sámi language as a result of active and/or passive suppression by dominant languages in their respective states of residence (Norway, Sweden, Finland, and Russia, e.g. Kuokkanen 2003). The Sámi media are thus faced with a choice: should they serve the whole community, which would mean switching to a majority language that more or less everyone understands? Or should they first and foremost try to preserve the languages that still are in use? In Finland, for example, the Sámi broadcaster is still very much a (North) Sámi language broadcaster, while services in Finnish are a secondary priority. In Norway, on the other hand, different media are used to serve different linguistic audiences: radio for the Sámi speakers and Internet for the Norwegian speakers (Markelin and Husband 2007). Similar trends prevail in Sweden, where the public service broadcasters want the broadcasting license to focus on media *on* the minority cultures, whereas the government insists on media *in* their languages (Moring 2017). Media for minorities thus have to balance many roles.

5. The impact of minority language media on the status and recognition of the language

An effect of language contact that is difficult to measure is its political impact, i.e. the impact, in terms of attitudes, that minority or indigenous media have on majority populations. For example, broadcasts in indigenous languages of sporting events in Wales or

Aotearoa, New Zealand, or Sámi television news broadcasts produced jointly by the Sámi television broadcasters, may reach hundreds of thousands of nonindigenous viewers. Although the impact of such broadcasts on the nonindigenous population is by no means comparable to the linguistic impact of majority media on the minority language community, it is likely that they improve the level of awareness. Such impact cannot be underestimated, as minority languages, to a large extent, rely on the goodwill of the speakers of majority languages.

In concrete terms, reaching nonspeakers may have specific effects: "Ōiwi TV of Hawaii for example has noted an increased use of correct diacritical marks after they have entered the public sphere" (Markelin 2017: 453). The use of <ā> in New Zealand English has also been influenced by the media, with a similar effect. Furthermore, according to Ladd (2007: Conclusion):

> [t]hough the Māori language experienced a substantial decline in the nineteenth century and in [the] first half of the twentieth century, educators, linguists and writers have contributed to Te Reo Māori's renewal. Māori's formal recognition by the NZ government was accompanied by programs that aimed to increase the number of Māori speakers, and by a more prominent place in the arts and in the media. You can now listen to the news in Māori, read bilingual government documents, and peruse one of several anthologies of Māori verse; you can even download a program to spell-check your Māori prose.

Similar developments have taken place elsewhere. The Sámi language technology center at Tromsø University (*Sámi giellateknologiija guovddáš*) works on language technologies to support the learning, teaching, and use of the language, for example, by developing tools for language checks, text analysis, speech and grammar support, and electronic dictionaries for six different Sámi languages (Aikio-Puoskari and Sámediggi/Sámegiela doaimmahat 2016).

The development of such language technologies for the least-spoken minority languages actualizes yet another language contact perspective: we cannot neglect the effects of minority media within the minority language communities. Pietikäinen and Kelly-Holmes (2011: 56) note that the Irish *Raidió na Gaeltachta* promotes the "heteroglossia of Irish in terms of the main dialects". Yet while these dialects may be promoted within a joint radio broadcasting service, the situation is more complex in the context of the Sámi languages, many of which are not mutually comprehensible. Just as majority media influence minority language communities, larger minority language communities also influence smaller ones. In terms of the Sámi media, most of the Sámi language content is produced in the most widely spoken variant: North Sámi. Smaller Sámi language groups receive far less media services and are further marginalized in the public sphere (Markelin, Husband, and Moring 2013). However, previous interviews with indigenous media workers from around the world showed that dying languages, too, have a role to play in the media. Markelin (2017: 453) notes that "the symbolic power of language(s) remain a central part of identity even when it is no longer transmitted to the next generation, and this shines through in the media policies of Indigenous broadcasters".

6. A new ecosystem and the trilingual trap

The digitalization of communication has changed the ecosystem of language use. While certain features have become easier, such as establishing media platforms online, the

digital realm has opened up a global communication marketplace that overthrows earlier regulatory policies. However, the market conditions for lesser used languages are crude: studies show a lag in the development of such media services online (Zabaleta et al. 2014).

Different types of media foster different patterns of use. Media use also varies according to social, age, gender, and regional differences. In developing media for minority languages, *substituting policies* (new media are expected to substitute old media) are a poorer choice than *additive policies* (new media are added to the supply; old media adjust their niche[s]). Minority language users are driven to majority language services if the services in their own language fall short on any particular platform, as they often have sufficient command of the majority language.

This richness of language skills leads to more varied media consumption among minority language users. Vincze and Moring (2013: 52) have shown that speakers of Swedish in Finland are divided into three substantial groups in their linguistic preferences for browsing on the web: main preference is Swedish (44 %), main preference is Finnish (32 %), and main preference is English (24 %). In contrast to this, speakers of the majority language (Finnish) prefer to browse the web in Finnish (92 %), with a small group preferring English (8 %). In a comparative study involving young Swedish speakers in Finland and young Hungarian speakers in Romania (16–17 years old), Vincze and Moring (2017) found similar patterns: a clear preference for browsing and using social media in English (over Swedish or Finnish) among the young in Finland, and evenly distributed preferences for both English and Hungarian when browsing and using social media among the young in Romania (over Romanian). Such trilingualism of minority language speakers may have far-reaching consequences for their language communities.

One of these is the tendency to migrate: between the years 2000 and 2015, almost 10 % of the Swedish population in Finland moved abroad (two thirds of them to Sweden). According to Kepsu (2016: 4), the return rate among movers to Sweden is around 50 %. Further, a considerable proportion of Indigenous people now live in urban spaces (Peters and Andersen 2013). This capacity to be 'all over the place' may be a positive factor for youth who have a rich language repertoire to fall back on. However, it may have negative consequences for the minority language community, as they are losing members of a linguistic community that is already under threat from such overwhelming language influx (most potently through the web).

7. Institutional completeness: a balancing act

The availability argument, i.e. the availability of media services in your own language, has become particularly relevant because of the current increase in globally accessible media content. The challenges faced by all smaller language communities are on the rise, as media of large language groups dominate the media landscape. This may challenge the *institutional completeness* of the (minority) media environment, and further erode *functional completeness* in media use (Moring 2007). Institutional completeness refers to access to media in a certain language, whereas functional completeness refers to the actual *use* patterns of services offered in the language. Few minority language media landscapes reach the level of institutional completeness that is achieved by Catalan and

Basque in Spain (Gifreu 2009), German in South Tyrol, Swedish in Finland (Moring and Godenhjelm 2011), or Russian in various states that were previously part of the Soviet empire. Upgrading the institutional status of minority languages in the media sector is essential, though not sufficient, for languages that wish to become more functionally complete.

Pietikäinen and Kelly-Holmes (2011) discuss the development of minority language media in relation to three (overlapping) eras in language policies: gifting, service, and performance (see Ricento 2006). *Gifting* refers to an era characterized by allocation of limited media space to a minority language by state-run media; the *service* era includes provision of a service to minority language speakers, with the goal of language revitalization; finally, media production in the *performance* era "explicitly makes use of the multilingual repertoires of the community and the target audience" (Pietikäinen and Kelly-Holmes 2011: 60).

According to Pietikäinen and Kelly-Holmes (2011: 63), "[…] the logic of the first era utilizes the norms of purity and monolingualism, the second welcomes also parallel multilingualism, and the third (to an extent) mixture and fluidity. This explains why, at the level of actually producing minority-language programmes, very different language policies may be adopted: from a majority- to a minority-language policy, and from parallel to heteroglossic multilingualism". Their study points to a new fluidity in an era where language contact is the rule rather than the exception, and where the media, too, adapt to a kind of 'hybridization' of cultures.

Recent studies show how vulnerable patterns of behavior related to language contact on the web may be a result of their contextual features. Johnson (2013) shows how fluent bilingual speakers of Welsh and English use their languages on Twitter. In the study, English messages slightly dominated the ones in Welsh (51 % versus 41 %), while less than 10 % of the messages were in a mixed language. However, the numbers slightly favored Welsh when tweets were in response to others. In fact, speakers preferred Welsh when the tweet was in response to other known biliterates (Johnson 2013: 108–109). However, Johnson (2013: 117) warns that "in terms of practical application, minority language users and supporters face challenges to discover methods in which Web 2.0 can work to the benefit of minority languages when they must compete against majority language resources, which, almost by definition, includes a greater number of speakers, resources for development of web portals, etc.".

On the basis of current developments in Irish, Sámi, and African languages, Kelly-Holmes (2014: 541) concludes that "[l]inked to a shift away from territorially-based speech communities […] has been an increasing commodification in relation to minority languages in the media. Here we can see a move from a rights based model to a lifestyle/consumption based model". Using the three language examples, she argues that the balance may tilt in favor of minority languages, as "new mediatized spaces create particular challenges and also opportunities for minority languages" (Kelly-Holmes 2014: 541) that may take different forms of hybridity and cultural demarcation. She concludes that "[p]erformance is a keyword that […] perhaps best illustrates the particular constellations of technology, agency, practice and ideology that we are currently experiencing" (Kelly-Holmes 2014: 539).

Graffman (2014: 13), in a study commissioned by the Scottish BBC Alba and based on 25 qualitative interviews with young (10–19 years old) speakers of Scottish Gaelic, remarked that "young people have developed their own strategies for keeping themselves

informed. Social media and peer-to-peer networks have great significance. It is often via their social networks that they obtain information, seeking further information on the basis of that. In Scotland, English is the main language when seeking information". On the web or on Skype, many of the interviewees also communicated in English with friends whom they knew to be Gaelic-speakers (Graffman 2014: 3).

Previously, due to spillover of the analog signals, broadcast services in minority languages of a particular state were also available in neighboring regions. Today, the digital space is confined within national borders for copyright reasons. Moring and Godenhjelm (2011) have shown how this problem particularly concerns minorities who speak the national languages of their smaller neighbor states (e.g. Danish in Germany, Swedish in Finland). The problem is not so significant for speakers of languages that are spoken by large populations with a global distribution of broadcasting (e.g. English, German, Russian, or French). Initiatives such as plans for a common digital space encompassing the entire EU could potentially have a favorable impact for all users of languages that are in official use in a neighboring state.

8. Discussion

The media, including increasingly web-based and social media, are today very much the locus of language contact. As we have seen, smaller languages are more vulnerable than larger ones in this reality. Therefore, we return to the question posed by Cormack (2007: 62), as quoted above: "In what ways can different media interact with other aspects of language use to contribute, directly or indirectly, to language maintenance in specific communities?"

Media studies show contradictory evidence about how media can contribute, given contextual conditions, to prevailing policies and resources available to linguistic communities. The digitalization of media has more or less revolutionized the spread of languages across borders and language groups. Studies show how we spend an increasing number of hours on one media device or another. Thus, the potential for language contact − be it with your own language, secondary, or tertiary languages − is considerable.

Research shows how language contact through the media is a double-edged sword for minority language communities. Young members of minority language communities − who are often tri- or even multilingual − can practice their language skills, for example, in English. Their multilingualism makes it easier for them to move, both domestically and internationally. Although rewarding at an individual level, this capacity may be problematic at a community level. Small language communities risk losing their vitality; the outside world may increase the vitality of a small community by providing new stimulus − but often at the cost of hybridization.

To help serve linguistic competence, minority language media need to be as functionally and institutionally complete as possible. Upgrading the institutional status of minority languages in the media sector is an essential, but not sufficient, policy measure for languages to become more functionally complete. Surveys among young people in minority language areas in Europe clearly indicate that many are abandoning their own language in favor of more widely spoken languages, be it the majority language or a *lingua franca* such as English. For minority language speakers, language contact through the media may thus be the equivalent of abandoning one's language.

Many media workers believe that their media institutions function as *language* institutions. Although clearly not the priority for all, maintenance and promotion of language is an important ambition for many media institutions and their employees. This is also the case in communities where a language is no longer widely used. Future research would do well to explore the interplay, over various media channels, between speakers of languages with different statuses.

Although the linguistic effects of this kind of language contact could be minimal (or nonexistent), the effects on attitudes and awareness could be all the more significant. Minority languages are in a disadvantaged position in relation to majority languages. Services to minorities are in many ways dependent upon the goodwill of the majority. Without awareness among majority language speakers about the existence and realities of minority language communities, the future of minority languages in the digital age is largely uncertain.

Presently, key issues are how digital language technologies, automatization in the production of media content, geo-blocking, and other aspects of digital humanities affect languages. These issues should be addressed, not just by linguists, but also by those engaged in research on language technology, as well as on the regulatory, economic, and political aspects of media. It is evident that the effort to form a viable cultural environment which is supportive of linguistic diversity and language contact is a multidisciplinary task, requiring efforts from researchers in many different fields.

9. References

Aikio-Puoskari, Ulla & Sámediggi/Sámegiela doaimmahat
2016 *Gullos sámegiella! Sámegielaid ealáskahttima buoremus vuogit ja našuvnnalaš politihka linjáť Suomas, Ruoťas ja Norggas*. Inari: The Finnish Sámi Parliament.

Androutsopoulos, Jannis
2014 Beyond 'media influence'. *Journal of Sociolinguistics* 18(2). 242–249.

Androutsopoulos, Jannis
2017 Media and language change: Expanding the framework. In Colleen Cotter & Daniel Perrin (eds.), *The Routledge handbook of language and media*, 403–423. London: Routledge.

Ball-Rokeach, Sandra & Melvin DeFleur
1976 A dependency model or mass-media effects. *Communication Research* 3. 3–21.

Blackwood, Robert
2013 French, language policy and new media. *Sociolinguistica* 27. 37–53.

Browne, Donald R. & Enrique Uribe-Jongbloed
2013 Introduction: Ethnic/linguistic minority media – What their history reveals, how scholars have studied them, and what we might ask next. In Elin Haf Gruffydd Jones & Enrique Uribe-Jongbloed (eds.), *Social media and minority languages: Convergence and the creative industries*, 1–28. Bristol: Multilingual Matters.

Cormack, Mike
2007 The media and language maintenance. In Mike Cormack & Niamh Hourigan (eds.), *Minority language media: Concepts, critiques, and cases studies*, 52–68. Clevedon: Multilingual Matters.

Cotter, Colleen, Daniel Perrin & Marlies Whitehouse
2017 Introduction. In Colleen Cotter & Daniel Perrin (eds.), *The Routledge handbook of language and media*, 1–4. London: Routledge.

Cottle, Simon (ed.)
2000 *Ethnic minorities and the media*. Buckingham: Open University Press.
Coupland, Nikolas
2010 Language, ideology, media and social change. *SPELL: Swiss Papers in English Language and Literature* 24. 127–151.
Coupland, Nikolas
2014 Language change, social change, sociolinguistic change: A meta-commentary. *Journal of Sociolinguistics* 18(2). 277–286.
Demont-Heinrich, Christof
2011 Cultural imperialism versus globalization of culture: Riding the structure-agency dialectic in global communication and media studies. *Sociology Compass* 5(8). 666–678.
Fishman, Joshua (ed.)
2001 *Can threatened languages be saved? Reversing language shift, revisited: A 21st century perspective.* Clevedon: Multilingual Matters.
Gifreu, Josep
2009 The Catalan communicative space: Still a strategic objective. *Catalan Journal of Communication and Cultural Studies* 1(1). 87–95.
Gitlin, Todd
1998 Public sphere or public sphericules? In Tamar Liebes & James Curran (eds.), *Media, ritual and identity*, 168–174. London: Routledge.
Graffman, Katarina
2014 Media behaviour among young Gaelic speakers: A comparative study in Scotland, Sweden, and Finland. Stockholm: Inculture (Mimeo).
Hall, Stuart
1997 *The local and the global: Globalization and ethnicity.* Minneapolis: University of Minnesota Press.
Hilgendorf, Suzanne K.
2013 Transnational media and the use of English: The case of cinema and motion picture titling practices in Germany. *Sociolinguistica* 27. 167–186.
Johnson, Ian
2013 Audience design and communication accommodation theory: Use of Twitter by Welsh-English biliterates. In Elin Haf Gruffydd Jones & Enrique Uribe-Jongbloed (eds.), *Social media and minority languages: Convergence and the creative industries*, 99–118. Bristol: Multilingual Matters.
Kelly-Holmes, Helen
2014 Commentary: Mediatized spaces for minoritized languages: Challenges and opportunities. In Jannis Androutsopoulos (ed.), *Mediatization and sociolinguistic change*, 539–543. Berlin: De Gruyter.
Kepsu, Kaisa
2016 Hjärnflykt eller inte? En analys av den svenskspråkiga flyttningen mellan Finland och Sverige 2000–2015 [Brain drain or not? An analysis of the Swedish language migration between Finland and Sweden 2000–2015]. *Magma-pamflett 2/2016*. http://magma.fi/post/2016/9/14/svenskspraki ga-nylanningar-toppar-flytten-till-sverige-magma-pamflett-2-2016. (accessed on 22 January 2019).
Kuokkanen, Rauna
2003 "Survivance" in Sami and First Nations boarding school narratives. *American Indian Quarterly* 27(3). 697–726.
Kuppens, An H.
2010 Incidental foreign language acquisition from media exposure. *Learning, Media and Technology* 35(1). 65–85.

Ladd, Heather
2007 English, Māori, and Māori English in New Zealand: A brief history of New Zealand. http://homes.chass.utoronto.ca/~cpercy/courses/eng6365-ladd.htm. (accessed on 10 April 2018).
Lehtola, Veli-Pekka
1997 *Saamelainen ääni: Saamen radio 1947−1997*. Helsinki: Yleisradio.
Markelin, Lia
2017 Indigenous voices in the public sphere: Analysis of approaches to journalism within the WITBN network. *Journal of Applied Journalism & Media Studies* 6(3). 443−461.
Markelin, Lia & Charles Husband
2007 The Sámi media, state broadcasting and transnational indigeneity. In Olga Guedes Bailey, Myria Georgiou & Ramaswami Harindranath (eds.), *Transnational lives and the media: Re-Imagining diasporas*, 115−132. Basingstoke: Palgrave-Macmillan.
Markelin, Lia, Charles Husband & Tom Moring
2013 Sámi media professionals and the role of language and identity. *Sociolinguistica* 27. 101−115.
Moring, Tom
2007 Functional completeness in minority language media. In Mike Cormack & Niamh Hourigan (eds.), *Minority language media: Concepts, critiques, and case studies*, 17−33. Clevedon: Multilingual Matters.
Moring, Tom
2017 *In* the language or *on* the language: Paradoxes in media policy for minorities. *Multiethnica* 36/37. 33−42.
Moring, Tom & Charles Husband
2007 The contribution of Swedish-language media in Finland to linguistic vitality. *International Journal of the Sociology of Language* 187/188. 75−101.
Moring, Tom & Sebastian Godenhjelm
2011 Broadcasting for minorities in big and small countries. In Gregory Ferrell Lowe & Christian S. Nissen (eds.), *Small among giants: Television broadcasting in smaller countries*, 179−202. Göteborg: Nordicom.
Moseley, Christopher (ed.)
2010 *Atlas of the world's languages in danger*, 3rd edn. Paris: UNESCO Publishing.
Ofcom
2014 *Techie teens are shaping how we communicate*. https://www.ofcom.org.uk/about-ofcom/latest/media/media-releases/2014/cmr-uk-2014. (accessed on 10 April 2018).
Peters, Evelyn & Chris Andersen (eds.)
2013 *Indigenous in the city: Contemporary identities and cultural innovation*. Seattle: University of Washington Press.
Pietikäinen, Sari
2008 Sami in the media: Questions of language vitality and cultural hybridization. *Journal of Multicultural Discourses* 3(1). 22−35.
Pietikäinen, Sari
2010 Sami language mobility: Scales and discourses of multilingualism in a polycentric environment. *International Journal of the Sociology of Languages* 202. 79−101.
Pietikäinen, Sari & Helen Kelly-Holmes
2011 Gifting, service and performance: Three eras in minority-language media policy and practice. *International Journal of Applied Linguistics* 21(1). 51−70.
Ricento, Thomas
2006 Language policy: Theory and practice − An introduction. In Thomas Ricento (ed.), *An introduction to language policy: Theory and method*, 10−23. Oxford: Blackwell.

Sayers, Dave
2014 The mediated innovation model: A framework for reaching media influence in language change (Focus article). *Journal of Sociolinguistics* 18(2). 185–212.
Stenberg-Siren, Jenny
2018 *Språk och språkideologier i radio och tv: Standardspråk och språkstandarder i finlandssvenska radio- och tv-nyheter* [Language and language ideologies in radio and television: Standard language and language standards in Finland-Swedish radio and television news] (Nordica Helsingiensia 50). Helsinki: University of Helsinki.
Tomlinson, John
2001 [1991] Cultural imperialism: A critical introduction. London: Continuum.
Vincze, Laszlo & Tom Moring
2013 Towards ethnolinguistic identity gratifications. In Elin Haf Gruffydd Jones & Enrique Uribe-Jongbloed (eds.), *Social media and minority languages: Convergence and the creative industries*, 47–57. Bristol: Multilingual Matters.
Vincze, Laszlo & Tom Moring
2017 Trilingual internet use, identity and acculturation among young minority language speakers: Some data from Transylvania and Finland. *European and Regional Studies* 12. 111–123.
Zabaleta, Iñaki, Carme Ferré-Pavia, Arantza Gutierrez, Itxaso Fernandez & Nikolas Xamardo
2014 European minority language media and journalism: Framing their marginal reality. *The International Communication Gazette* 76(3). 275–295.

Tom Moring, Helsinki (Finland)
Lia Markelin, Helsinki (Finland)

62. The politics perspective on language contact

1. Introduction
2. State formation, language, and collective identity
3. Linguistic identity and the politics of recognition
4. Complex diversity and the new multilingualism
5. The multilingual citizen: opening linguistic gates and securing linguistic ties
6. References

1. Introduction

When dealing with language contact, social science approaches in general, and political science approaches in particular, tend to adopt a perspective under which individual choices regarding language use do not so much reflect spontaneous and innovative agency, but rather the impact of power relationships, be it at the macrolevel (e.g. the state), the mesolevel (e.g. an educational institution), or the microlevel (e.g. the family). Such relationships tend to be characterized by unequal access to resources and by situations of domination. To make this point clearer, let us briefly recall two of the most prominent definitions of power offered by classical political sociology and political science.

https://doi.org/10.1515/9783110435351-062

According to Max Weber's quasi-canonical definition, power "is the probability that one actor within a social relationship will be in a position to carry out his own will despite resistance, regardless of the basis on which this probability rests" (Weber [1968] 1978: 53). An interesting alternate, and perhaps more poignant, definition is given by Karl W. Deutsch (1963: 111), according to whom power "*is the ability to afford not to learn* [emphasis in the original]". To remark that both definitions have a linguistic dimension can hardly be considered an over-interpretation. This is so even in a 'simple' monolingual setting, as exercising power will in most cases require linguistic mediation. The politics of language tends to become much more salient, however, in any multilingual context, i.e. in a context of language contact, which in the modern world has become a virtually ubiquitous phenomenon. Thus, following Weber's (1978) definition, it seems obvious that a German Junker would impose his will *vis-à-vis* his Polish serf in East Elbia around 1900 by addressing him in German, thereby showing, if only implicitly, that a relationship of quasi-feudal oppression converges with a relationship of national and linguistic domination. The linguistic factor is even more palpable in the definition of Deutsch (1963), as 'affording not to learn' is a resource unevenly distributed across the members of linguistic groups, both nationally and transnationally. It thus seems more 'affordable' for an English-speaking banker in California not to learn Spanish, than for a Spanish-speaking plumber in the same state not to learn English. In a similar way, anglophone financial managers in Frankfurt may continue speaking only English in their professional activities, whereas normal staff members in the same business will in all likelihood be required to prove that they have fluency in German, in addition to some level of English proficiency.

As these two examples make evident, from a political perspective, issues of language contact inevitably become intermingled with issues of, and conflicts over, status attribution. In general terms, such status attribution has absolutely nothing to do with the internal features of a language. For instance, contrary to a frequent misunderstanding, the fact that English has attained the quality of a trans-European *lingua franca* is not at all connected to it having a grammatical structure that may seem more accessible to nonnative speakers than the structure of, say, Finnish or German. It instead reflects a situation in which the economic, political, and cultural status of English at the global level has an impact that has become more and more palpable at the European level as well, albeit at the expense of other regional *linguae francae*, such as French, German, or Russian.

Scholars studying contact from the perspective of linguistics advocate for establishing "a number of theoretical principles through which contact can be interpreted and appreciated in an integrated manner" (Matras 2009: 2). However, when it comes to interpreting contact-related change, they tend to prioritize "innovations that individual multilingual speakers introduce into discourse in a multilingual setting" (Matras 2009: 5). In contrast, the approach adopted in this contribution instead emphasizes the structural factors limiting and/or enabling individual agency. Bringing politics into focus may thus not only be a worthwhile effort in terms of interdisciplinarity, but also be a key element for properly understanding what is at work when modern linguistic identities are created, reproduced, and changed in a multilingual environment. In this respect, the almost complete absence of the political perspective in recent standard publications in the field of contact linguistics is striking; in addition to the above-quoted book by Matras (2009), see e.g. Hickey (2010) and Winford (2003).

In short, what a linguist might interpret as *the political dimensions of language contact*, a political scientist could very well reinterpret as *the linguistic dimensions of "political contact"*. To offer an assessment of how politics has been, and continues to be, a constitutive element in the linguistic makeup of modern societies, this contribution proceeds in four steps. In Section 2, I show how language has been a key element in the dynamic of modern state-making and nation-building. Section 3 deals with the issue of recognition in settings where minorities mobilize for the purpose of overcoming status inequalities in a multilingual setting. Section 4 revolves around the implications of complex diversity for the articulation of a new politics of multilingualism. Finally, in Section 5, I sketch out some normative conclusions regarding how the ongoing transformation of diversity patterns should entail the generalization of multilingual repertoires in multilingual societies.

2. State formation, language, and collective identity

Modern history has largely been the history of the development of the discrete political units we call nation-states. By and large, the structure of these units reflected the converging paths of the territorial integration of societies with the standardization of these societies along cultural and linguistic lines. The imprint of this dynamic was particularly strong in the context of Europe, where the processes of state construction led to an increasing overlap of cultural and linguistic borders. This overlap was obviously not the result of the free and spontaneous interaction of different language communities, but reflected deliberate political strategies. In combination with religion, language became a crucial factor in the making of national identities. In historical retrospective, the Reformation can indeed be assigned a key role in initiating the processes of establishing national languages, in particular in the northern and central regions of Europe. On the one hand, the borders between areas of rule came to reflect the different religious observances of the rulers. On the other hand, the rupture with the Holy See led to the making of a more or less explicit bond between national churches and national languages in those areas controlled by Protestants. This bond became particularly powerful due to the "coalition between Protestantism and print-capitalism" described by Anderson (2016: 40), a coalition that facilitated the steady spread of a particular language variety or language among the population of a particular territory. Thus, in Scandinavia one can already detect at an early stage an intertwining of linguistic and religious affiliations in the construction of the Nordic peoples' identities. With the establishment of Lutheran state churches, vernacular languages became standardized both as a means for conveying the religious liturgy and as templates of mass education in the course of the successive extension of schooling. In the emerging system of Westphalian states, the principle of territorial sovereignty became synonymous with the forging of religiously and linguistically homogeneous domains of rule. *Cuius regio, eius religio*, the motto of the Peace of Augsburg (1555), the first confessional settlement after the beginning of the Reformation, thereby had its twin companion in the principle *cuius regio, eius lingua*.

The interplay of nation-building and linguistic standardization has been systematically analyzed in the work of the Norwegian political sociologist Stein Rokkan (1999). According to his model of political development, European modernity was largely contin-

gent upon a "proto-nationalist" dynamic that placed language − and indeed, in most cases, exclusively *one* language − in a prominent position as an identity marker of a territory, a marker creating ever closer links between rulers and their subjects. In this sense, we can speak of an initial predominance of a *top-down* logic in the processes of linguistic standardization, the main goals of which were achieving administrative efficiency through the introduction of uniform linguistic standards (the French *Ordonnance de Villers-Cotterêts* [1539] was an early manifestation of this purpose; Jacob and Gordon 1985: 111−112), and at a somewhat later stage, securing common patterns of loyalty among the population of a given territory by making this territory culturally homogeneous. As a result of the fusion of democratic and national ideas observable all over Europe (and the world) in the aftermath of the French Revolution, the top-down dynamic was gradually reshaped, if not replaced, by *bottom-up* mobilizations. For the French Jacobins, French was no longer the language of dynastic rule, but rather an icon of national emancipation (Balibar and Laporte 1974). What in France are now called 'regional languages', such as Breton and Basque, were in contrast considered obsolete symbols serving the antirepublican cause. In the 19th century, bottom-up language nationalism would find many followers in those parts of the Continent under Austro-Hungarian, Russian, and Turkish control. Here, nation-building did not follow the rhythm of a historical *longue durée*, and instead went hand in hand with the breakdown (rather than with the reform) of former empires, so that the idea of the national language turned out to be more politically explosive than in most North and West European countries. Thus, the struggles directed against the Ottomans or the Habsburgs quickly merged with linguistic nationalisms that saw their very *raison d'être* in substituting the imperial 'high' language with the oppressed vernaculars (Kraus 2016: 42−43).

A key motif in the processes of linguistic mobilization at the high time of nationalism in Europe was the conviction that the soul of a particular nation found its utmost expression in its particular language. For Czechs and Slovenians, for Latvians and Finns, their respective languages embodied the same dignity as the languages of supposedly 'higher' cultures did. The defense of the vernaculars thereby became a central battlefield in the quest for national emancipation (Hroch 1985). In this quest, the patterns of collective identification became manifest through the attachment of one people to their singular national language. This view is succinctly captured in the approach of the philonationalist historian John Wuorinen (1931: 53), who writes about how (parts of) the Swedish-speaking upper classes in Finland chose "to adopt Finnish as their mother tongue" to ensure that the Finns were able to "become a fully united nation". As a result of the wave of national-*cum*-linguistic mobilizations that stretched from the 'spring of nations' in 1848 through the 20th century, many groups had the possibility to establish 'their' standard against the previously dominant one, thereby turning former majorities into postimperial minorities.

In the long run, the institutional framing of linguistic identities in Europe led to the situation we widely experience today. The 'standard' European citizen is the European whose standard language is the majority standard. At the same time, and with very few exceptions, this majority standard is directly deducible from the official denomination carried by the nation-state to which this citizen belongs, be it Bulgarian (i.e. not Turkish) in Bulgaria, Italian (i.e. not German) in Italy, or Norwegian (i.e. not Sámi) in Norway. As such, what Europeans tend to regard as the language which defines the nation and its citizens is, in the first place, the majority language.

3. Linguistic identity and the politics of recognition

In light of the macrohistorical dynamic described in the previous section, it seems quite plausible to view modern states as linguistic assimilation machines. As pointed out by Calvet (1974), "glossophagia", or the devouring of languages through the institutionally enforced adoption of a common standard, is closely related to the structures of political domination in modern societies. The banning of what would become, at best, minority languages from central domains of public life was a feature not only of European polities, but also of many states in most other parts of the world which were eager to follow the European example.

As seems to be indicated by the steady decline of the number of languages spoken on our planet, glossophagia is a persistent phenomenon. This does not mean, however, that it is an uncontested one. Many people are apparently prepared to make significant efforts to maintain what they see as 'their' language, however 'small' that language may be. These efforts become especially salient in the case of dying or dead languages, which are revitalized by minority activists who consider them to be pivotal symbols of their identity. While cultural standardization has been a key feature in the creation of the modern world, its success has thus far been effectively limited by the quest to maintain cultural and linguistic authenticity, a quest that seems to be as modern as the push for homogeneity, albeit in a different way (May 2012: 173−174). This is best explained by a concept which the Canadian political philosopher Charles Taylor (1992) has aptly coined, namely 'the politics of recognition'. Following a path paved by representatives of the German enlightenment, such as Johann Gottfried Herder and Wilhelm von Humboldt, a language is not only a resource we need for purposes of instrumental communication, but also the expression of a particular life-world connected to a specific cultural community. Constituting an irreducible social web, language links our dignity as individuals to the status of this community. If we are to enjoy equal dignity as citizens of a democratic society, the collective web we draw on to build our individual identity must be protected by institutional provisions that permit its reproduction. In other words, the quest for authenticity becomes inextricably interwoven with demands for self-determination.

The politics of linguistic recognition thereby typically aims to achieve an equal linguistic status − or equal dignity (see Van Parijs [2011: 117−132] on the politics of language and the issue of equal dignity) − for one's 'own' language, i.e. the language one feels expressively attached to. A minority group mobilizes on this basis, with the claims that its self-categorization be accepted by the majority and that it be given the right to use its language in functional domains hitherto restricted to the majority language. As cases such as the South-Tyroleans in Italy or the Swedish-speaking Finns show, the rationale of such demands is not always linked to situations of socioeconomic discrimination, rather what is ultimately at stake in the dynamic of linguistic mobilization in Western societies seems to be the demand for affirmation of a distinctive cultural identity, language being the main icon of this identity (Allardt 1980).

Against this backdrop, a critical step for achieving a comprehensive understanding of group conflicts in multilingual contexts is to establish a connection between the role of recognition in language policy and the analysis of diglossia (Kraus 2008: 101−102). In the work of Joshua Fishman (1967, 1971), *diglossia* describes a multilingual setting within a given society in which two or more languages are assigned clearly separate

functional domains. The political relevance of diglossia stems from the hierarchical ranking frequently associated with this separation: a 'high' language occupies the upper domains, such as administration, education, and the financial world, while the 'lower' language(s) are restricted to use in private and/or informal situations. The tearing down of such hierarchies has been very much at the heart of conflicts over linguistic recognition in the West and elsewhere. The interplay of diglossia and bilingualism allows for a variety of combinations, and the dynamic of this interplay tends to indicate major sociopolitical transformations. The historical evidence to hand offers some hints that non-diglossic bilingualism tends to be unstable (Laponce 1987). Ensuring that the equal coexistence of two (or more) official languages within the same political unit does not entail struggles over recognition (in terms of the equal dignity of the language communities) obviously requires a great amount of institutional creativity, as the cost of achieving and maintaining a bilingual repertoire should not be mainly carried by one language group while sparing any expenditure from the other. Ideally, linguistic recognition should mean that members of a minority have sufficient opportunities to lead an everyday life that, with respect to the presence of their 'own' language, is similar to that to which members of the majority are accustomed. This always implies a substantial, politically induced expansion of the range of functional domains in which the minority language is used. Eliminating the collective resentment caused by hierarchies and asymmetries that speakers of different languages may be exposed to despite the officially equal status of their language *vis-à-vis* another language is one of the substantive challenges of a democratic and equitable language policy. At any rate, this is the lesson to be learned from political developments in multilingual societies such as Belgium, Québec, and Catalonia over the last few decades.

4. Complex diversity and the new multilingualism

To this point, my assessment of the politics of recognition has focused on conflicts between majorities and 'old' minorities. In its most general terms, in Europe, the concept of *majority* refers to a group whose identity coincides with the titular nation established in the process of state construction (such as the Romanians in Romania, or the Swedes in Sweden). 'Old' minorities, in contrast, may be seen as victims of this dynamic. They are constituted by national and/or indigenous minority groups who were not included in the 'we' officialized in the making of particular nation-states (as is the case with the Turks in Bulgaria, the Basques in France and Spain, and the Sámi in Northern Europe, to name only a few examples). The historical grievances suffered by these groups became the very basis of their political demands for autonomy and language rights. In addition, the patterns of cultural belonging characteristic of European modernity have been substantially affected over recent decades by the emergence of a third layer of diversity: 'new' minorities formed by immigrants, and their first-, second-, and third-generation descendants (be they people with a South Asian background in Britain, 'German Turks', or French citizens of Maghrebin origin).

As we will see, classifying people along these lines still implies operating within the categories of a 'simple' diversity, according to which nation-states and societies operate on the basis of neatly drawn demarcations between the layers of diversity they

have incorporated. This normally goes hand in hand with assigning individuals and groups to specific 'identity slots' in correspondence with institutional arrangements that make for a ranked system of recognition (Kraus 2012: 9–13). When it comes to language, the rights of the majority thereby apply evenly across the entire territory of the state: Spanish, for instance, has official status throughout Spain. 'Old' minorities may benefit from provisions that acknowledge their linguistic existence in their homeland under some form of regional autonomy regime: in the Spanish Basque Country, Basque is co-official with Spanish. In contrast, to the extent they exist at all, the cultural rights granted to 'new' minorities are not meant to allow them to maintain a particular group identity, but rather to facilitate their social and economic integration into the host society; thus, mother tongue instruction for the children of immigrants is still an exceptional phenomenon in most European countries (see Banting et al. [2006] for a systematic overview of the range of multiculturalism policies in North America, Australia, and several West European countries).

The problem with a hierarchically layered approach of this kind is that it entails a static view of identity formation, conceiving of the three layers of diversity as if they corresponded to distinct, clearly separable frames of collective belonging. Introducing the concept of 'complex' diversity challenges this view. In contrast to the notion of a 'simple' diversity, it does not aim solely to emphasize that our societies have become more diverse everywhere due to migration flows and the transnationalization of an increasing number of cultural domains, nor is calling diversity 'complex' simply a rhetorical device designed to stress that we have become much more aware of the value of diversity than in the times of high modernity. The key point of the concept of complex diversity is that it underscores the multidimensionality and fluidity of diversity today. Complex diversity characterizes constellations in which cultural identities and social cleavages overlap and intermingle in changing and dynamic ways. Not only has the incorporation of new layers of cultural differentiation made our societies more diverse, but the very layers (or building blocks) of diversity are themselves becoming increasingly porous and heterogeneous. Thus, to give a concrete example, a Spanish citizen of North African origin in Barcelona may well be a Muslim, speak Catalan and Arabic (or Amazigh) in addition to Spanish, and be supportive of the cause of Catalan independence. To file this person under one of the previous three 'simple' identity categories does not look like a highly productive sociological exercise. This observation is also obviously relevant when it comes to understanding this person's linguistic affiliations and loyalties.

Following the logic of this example, if we apply the concept of complex diversity to the politics of language and language contact, the most promising settings to examine (in terms of the heuristic possibilities they offer) are the almost laboratory-like constellations where the patterns of the old elements constitutive of autochthonous linguistic majority and minority groups confront a vast spectrum of new elements of linguistic heterogeneity. Such constellations are characterized by the parallel use and intertwining of autochthonous languages (with majority and minority status), immigrant languages, and a *lingua franca* (which sometimes may be identical to one of the autochthonous or immigrant languages). Such constellations reflect the proliferation of new forms of mobility. These trigger a transnationalizing dynamic that decouples the territorial bases of identity construction from actual cultural practices. In this respect, many urban areas in Europe are coming close to sociolinguistic configurations that seem to resuscitate the

Continent's medieval past. The risk involved with these configurations is that they re-install structures that allow for the overlap of linguistic differentiation and stratification, thereby creating segments that remain largely isolated from each other and mutually reinforcing the cultural and social division of labor, as was typically the case in medieval urban contexts (Geary 2001: 40). In this kind of scenery, in a city such as Brussels, we would find lobbyists and IT staff relying exclusively on English, Moroccan or Turkish immigrants working in low positions and leading their lives largely in Arabic or Turkish, and the local middle classes (with a French speaking majority and a Dutch speaking minority) 'sandwiched' between the upper and the lower segments. This would make our near future strikingly − and problematically − similar to our medieval past. On the other hand, the rise of complexity in European cities may give rise to new forms of linguistic identity-building that could overcome cultural segmentation through acquisition of functional multilingual repertoires by the bulk of a city's population. In this respect, the 'new' medievalism would break in appealing ways with the legacies of the old Middle Age.

To evaluate whether, and in what ways, complex diversity is reframing the relationship between identity-building and the politics of language is ultimately an empirical question. To address this question in illuminating ways, the best candidates to study do not seem to be the usual suspects, i.e. 'global capitals' recurrently chosen when it comes to assessing the social and political impact of 'super-diversity', such as Paris or London (Vertovec 2007). In accordance of what has been argued in this section, the more promising cases for studying the effects of the new diversity on multilingualism are urban areas where the interplay of historically inherited patterns of multilingualism with new lines of linguistic differentiation introduced by immigration leads to remarkably rich forms of complexity. Thus, cities such as Brussels, mentioned above, as well as Barcelona, Helsinki, Luxembourg City, Riga, and Tallinn could be considered particularly apt locations to assess the new politics of language (see Kraus [2011] for a comparative attempt in this direction).

5. The multilingual citizen: opening linguistic gates and securing linguistic ties

To point to a trend towards a new multilingualism in important European conurbations does not necessarily imply that we can assume the days of the traditional 'one language, one state' formula are numbered, and that monolingual majority citizens are doomed to lose their privileges to ascendant mobile polyglots. The blunt truth is that in Europe (as in other parts of the globe), official monolingualism is still the standard. Thus, in the European Union (EU) as it stands today (still including the United Kingdom at the time of writing), only 4 out of 28 states have a clear-cut multilingual profile. This is a figure that makes for only 14 % of EU member states. That proportion would be even lower if it were indicated in terms of population, as the four countries − Belgium, Finland, Ireland, and Luxembourg − all belong to the group of demographically smaller states in the Union.

As I have argued in this chapter, the reluctance of modern states to 'normalize' language contact through the adoption of open multilingual formulas is intrinsically linked

to issues of political authority and domination. It is true that, even under conditions of official state monolingualism, many European states have in recent years adopted regulations that afford to members of 'old' minorities a minimum of institutional protection, typically by granting them linguistic rights at the personal level or by establishing linguistic autonomy provisions that apply in some parts of the state territory (i.e. at the regional level). Concessions of this kind (which have found their most prominent international codification thus far in the *Charter for Regional or Minority Languages* elaborated by the Council of Europe in 1992) bear witness to contemporary constitutionalism's increasing reluctance to consider compulsory assimilation to be a legitimate means for achieving political stability. On the contrary, safeguarding the cultural rights of minorities has become an important condition that a state must meet if it wants to be viewed as a consolidated liberal democracy. Nonetheless, this approach toward minority protection should not be mistaken for a clean break with the legacies of a political modernity largely shaped by its push for homogenization, the main result of which is the Westphalian order of states that has been structuring Europe's politics since early modern times (for an elaboration of this point, see Kraus 2018: 11−16).

Minority provisions and regional autonomy are intended to attenuate the assimilationist pressures to which minorities are structurally exposed. Yet they are not supposed to meet this objective by challenging the foundations of a state system constructed by − or, rather, in the name of − dominant majorities. Ultimately, their intention is to stabilize the structures of established statehood by avoiding exacerbation of minority conflicts. Minority rights would not be labelled as such if they equipped the members of subordinate linguistic groups with the material levels of self-determination that those citizens who belong to the majority enjoy. Ultimately, defining a politically minoritized collectivity as 'protected' or 'autonomous' is synonymous to saying that it is not sovereign. For the minority in question, receiving some type of recognition will still be preferable to receiving no recognition at all, yet we are still speaking of a recognition that is inextricably tied to the authority of a given state to define (as well as, ultimately, withdraw) the very mechanisms which formalize the recognition of said linguistic minority. The power relations underlying this procedure have an unavoidably asymmetric, if not unequal, character. In the context of majority-minority politics, concepts such as 'language contact' and 'language rights' can almost be seen as euphemisms for various forms of cultural domination. In situations of complex diversity, coping with the challenges of acquiring a multilingual repertoire and developing reflective identities therefore typically remains the business of those citizens who belong to a minority group, who have to make do with multiple − and sometimes contradicting − identity options and language choices. For a minority citizen, coping with diversity is an indispensable task for as long as they reject assimilation. In contrast, for a majority citizen, diversity is an optional offer which they may embrace − or not.

There seems to be a sharp discrepancy between the trends we observe empirically 'on the ground' in the multilingual cities just mentioned, and what continues to be the dominantly monolingual institutional profile of European nation-states. However, the challenges of a new multilingualism, which up to now may have largely been considered a predominantly urban phenomenon, should be seriously taken into account at the national and transnational levels as well. It is clear that with regard to language (as presumably in other ways as well), complex diversity involves difficult individual choices. To return to the example of the citizen of Barcelona of North African origin: let us assume that

this citizen is the child of Moroccan immigrants with a Berber background. To maintain a linguistic repertoire that, on the one hand, pays tribute to the culture of their grandparents, and on the other hand, is prepared to fully meet the linguistic requirements of a multinational and globalized environment, they would have to learn Amazigh and Arabic (which are both offered by some local schools, at least as an option), as well as Catalan, Spanish, and obviously, English. This makes for proficiency in five different languages. To acquire such skills − even at strongly varying competence levels − is certainly not a minor task to fulfil. At the same time, fostering a habitat that offers our Berber-Catalan (with a Spanish passport and relatives in Morocco) the facilities to fully realize their multilingual potential calls for an educational infrastructure that is lightyears away from what appears feasible in the classrooms of most European schools at present, even more so if the multilingual curriculum to be implemented is built on equitable and nonhierarchical criteria.

So, the stakes are certainly high. They are nonetheless well worth pursuing, if the polities and citizens of the 21st century are to supersede the monolingual bias characteristic of the period of national rule. Cities, regions, and ultimately also states will be more and more often required to develop responsive approaches in order to tackle complex linguistic diversity in schools and public organizations. This is not only the case because knowledge-based societies must view the multilingualism of their citizens as a valuable economic asset which enhances cognitive mobility; much more important than the economic payoffs is that multilingualism is a key resource for fostering the intercultural understanding on which complexly diverse societies rely to combat ethnocultural ghettoization as well as the injustices of a polarized and hierarchic diglossia. From the point of view that acknowledges the linguistic embeddedness of civic identities, the worth of languages stretches far beyond their strictly instrumental functions. In particular, cities composed of an increasingly diverse citizenry should allow for a set of variable multilingual repertoires in designing the profile of their institutions. Adopting an active approach to "citizenization" (Tully 2008: 311) is a basic step toward creating structures of an effective political participation in complexly diverse societies. This may well imply the need for questioning the idea that there is one standard curriculum which targets the needs of all citizens in an equitable way, regardless of their differing sociocultural backgrounds.

Accordingly, conceiving of multilingualism as a central component of 'citizenization' under conditions of diversity should not be confused with adopting a 'one-size-fits-all' formula that is equally applicable in, say, Riga and Barcelona, or Amsterdam and Berlin. Nevertheless, the ideas sketched out in this section convey one important normative implication, namely that in our societies, all institutional approaches towards linguistic diversity should contribute to generalizing multilingual repertoires in ways that allow people to open linguistic gates and to secure linguistic ties. These repertoires should then be tailored in accordance with the specificities of the sociopolitical and sociocultural context where they are to be implemented. To speak of linguistic gates and linguistic ties takes up classical conceptualizations of an antinomy that seems inherent to our existence in modernity, an antinomy that has been labeled in terms of: mechanic versus organic solidarity (Durkheim), community versus society (Tönnies), and life-world versus system (Habermas). Against the contextual background of such theories, the German-British sociologist and political scientist Ralf Dahrendorf (1979: 30−33) proposed the terms 'options' and 'ligatures' to identify the two mechanisms he regarded as constitutive

of the life opportunities individuals have in contemporary societies. For Dahrendorf, options and ligatures do not necessarily stand in stark opposition to each other, but rather may be seen as key elements of a complementary, and potentially fertile, relationship. If we transfer this approach to the realm of linguistic diversity, it seems obvious that languages-as-options open the gates to a universal *koiné*, to a virtually unlimited realm of communication; languages-as-ligatures, in contrast, are the ties that link individuals to a bounded community built upon institutionalized collective practices to which they feel attached (Kraus and Kazlauskaitė-Gürbüz 2014: 521–523). As can be easily deduced from the examples of multilingualism discussed in this contribution, there is a tension between the mechanisms that create linguistic options, and those that uphold linguistic ligatures, even if we assume *à la* Dahrendorf, that the tension does not rule out complementarity. In the case of our Berber-Catalan citizen, for example, the more they focus on the options side, the more they may prioritize Spanish, Arabic, and/or English, which are all widespread *linguae francae*, at the expense of Catalan and Berber, the communicative reach of which is more localized, however strong their emotional spell at the local level may be.

With this example in mind, it should be clear that the approaches advocated in this contribution are those approaches that turn painful trade-off situations into positive-sum games. Languages as options that provide opportunities for choice must not be weighed against languages as ligatures that create social bonds. At the individual level, the critical advantage languages offer in comparison to religions is that they can be added to a feasible multilingual repertoire, whereas becoming 'multireligious' is a less realistic possibility. In this respect, the multilingual citizen should be seen as both a genuine subject and a desirable strategic object of politics in settings which are becoming more and more diverse in complex ways. From the perspective adopted in this chapter, institutionalizing multilingual citizenship is a key challenge for a politics that is up-to-date and prepared to overcome the rigidities of a modernity that remains ensnared in one-dimensional cultural imageries (Toulmin 1992). It should also be clear that coming to grips with this challenge will require defining more comprehensive and context-sensitive linguistic repertoires for the citizens of complex societies. Managing this challenge is not just a question to be tackled in terms of innovating educational curricula and facilitating language acquisition. It is a question that ultimately points to how we want to conceive of equality and fairness in contemporary polities, with the goal of finding ways to overcome the unjust hierarchies baked into the Westphalian state system, and thereby developing and implementing new notions of linguistic sovereignty.

6. References

Allardt, Erik
1980 Prerequisites and consequences of ethnic mobilization in modern society. *Scandinavian Political Studies* 3(1). 1–20.

Anderson, Benedict
2016 [1983] *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.

Balibar, Renée & Dominique Laporte
1974 *Le français national*. Paris: Hachette.

Banting, Keith, Richard Johnston, Will Kymlicka & Stuart Soroka
2006 Do multiculturalism policies erode the welfare state? An empirical analysis. In Keith Banting & Will Kymlicka (eds.), *Multiculturalism and the welfare state: Recognition and redistribution in contemporary democracies*, 49–90. Oxford: Oxford University Press.

Calvet, Louis-Jean
1974 *Linguistique et colonialisme: Petit traité de glossophagie*. Paris: Payot.

Dahrendorf, Ralf
1979 *Life chances: Approaches to social and political theory.* Chicago: University of Chicago Press.

Deutsch, Karl W.
1963 *The nerves of government: Models of political communication and control*. New York: Free Press.

Fishman, Joshua A.
1967 Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues* 23(2). 29–38.

Fishman, Joshua A.
1971 The sociology of language: An interdisciplinary social science approach to language in society. In Joshua A. Fishman (ed.), *Advances in the sociology of language*, vol. 1, 217–404. The Hague: Mouton.

Geary, Patrick J.
2001 *The myth of nations: The medieval origins of Europe*. Princeton, NJ: Princeton University Press.

Hickey, Raymond (ed.)
2010 *The handbook of language contact*. Chichester: Wiley-Blackwell.

Hroch, Miroslav
1985 *Social preconditions of national revival in Europe: A comparative analysis of patriotic groups among the smaller European nations*. Cambridge: Cambridge University Press.

Jacob, James E. & David C. Gordon
1985 Language policy in France. In William R. Beer & James E. Jacob (eds.), *Language policy and national unity*, 106–133. Totowa, NJ: Rowman & Allanheld.

Kraus, Peter A.
2008 *A Union of diversity: Language, identity and polity-building in Europe*. Cambridge: Cambridge University Press.

Kraus, Peter A.
2011 The multilingual city: The cases of Helsinki and Barcelona. *Nordic Journal of Migration Research* 1(1). 25–36.

Kraus, Peter A.
2012 The politics of complex diversity: A European perspective. *Ethnicities* 12(1). 3–25.

Kraus, Peter A.
2016 The politics of linguistic identity in Europe: Between the expression of power and the power of expressivity. In Christina Späti (ed.), *Language and identity politics*, 38–57. New York: Berghahn.

Kraus, Peter A.
2018 Between minority protection and linguistic sovereignty. *Revista de Llengua i Dret / Journal of Language and Law* 69. 6–17.

Kraus, Peter A. & Rūta Kazlauskaitė-Gürbüz
2014 Addressing linguistic diversity in the European Union: Strategies and dilemmas. *Ethnicities* 14(4). 517–538.

Laponce, Jean A.
1987 *Languages and their territories.* Toronto: University of Toronto Press.

Matras, Yaron
2009 *Language contact*. Cambridge: Cambridge University Press.
May, Stephen
2012 *Language and minority rights: Ethnicity, nationalism and the politics of language*, 2nd edn. London: Routledge.
Rokkan, Stein
1999 *State formation, nation-building, and mass politics in Europe*. Oxford: Oxford University Press.
Taylor, Charles
1992 *Multiculturalism and "the politics of recognition"*. Princeton, NJ: Princeton University Press.
Toulmin, Stephen
1992 *Cosmopolis: The hidden agenda of modernity.* Chicago: University of Chicago Press.
Tully, James
2008 *Public philosophy in a new key: Volume I: Democracy and civic freedom.* Cambridge: Cambridge University Press.
Van Parijs, Philippe
2011 *Linguistic justice for Europe and the world*. Oxford: Oxford University Press.
Vertovec, Steven
2007 Super-diversity and its implications. *Ethnic and Racial Studies* 30(6). 1024–1054.
Weber, Max
1978 [1968] *Economy and society* [eds. Guenther Roth & Claus Wittich]. Berkeley: University of California Press.
Winford, Donald
2003 *An introduction to contact linguistics*. Oxford: Blackwell.
Wuorinen, John
1931 *Nationalism in modern Finland*. New York: Columbia University Press.

Peter A. Kraus, Augsburg (Germany)

63. Social psychology

1. Introduction
2. Language as a symbol of social identity
3. Language and intergroup context: a social psychological approach
4. Conclusion
5. References

1. Introduction

Humans are social animals. To fulfill their need for positive relationships with fellow humans and to facilitate mutual understanding, they developed an important communicative tool: a shared verbal language. Language can be broadly defined as a system of communications wherein nonsense utterances combine and make sense by virtue of shared rules existing between communicating individuals (Robinson 2001). This defini-

https://doi.org/10.1515/9783110435351-063

tion of language is certainly oversimplified, yet its message is clear: language expresses meaning to those who hear it (Chomsky 1957). For this reason, language and language behaviors (vocabulary, intonations, enunciations, and accents) facilitate mutual understanding when individuals communicate with each other in an interpersonal context. It is thanks to common language behaviors that a professor can ask her colleagues to take a seat, that a policeman and a citizen can interact normally during a traffic stop, and that an individual can obtain a cup of coffee in a coffee shop.

However, language behavior conveys information that extends far beyond the interpersonal context. The language behavior used can also carry powerful social and intergroup messages, depending on the specific context in which language is used. A professor will speak differently to a family member than to her colleagues, even when asking both to take a seat. A white policeman may address a white man differently than a black man during a routine traffic stop. In the early 1990s in areas of Bosnia and Herzegovina controlled by Croatians, ordering a coffee with a Croatian pronunciation (*kava*) would get the person a decent cup of coffee, while ordering coffee with a Bosnian-Muslim pronunciation (*kahfa*) could instead be met with the wrong end of a gun (Dragojevic, Gasiorek, and Giles 2016: 9). These examples illustrate how interpersonal communications are often framed in larger social and intergroup contexts. These intergroup contexts impact how individuals communicate with each other. In the case of the professor and the policeman, their intonations and enunciations depend on the professional or racial membership of their audience, respectively, while the client in the coffee shop might adopt an accent that is more acceptable in the context of ethnic conflict. For any of these individuals, using the wrong language behavior at the wrong moment can have serious consequences, ranging from being socially ostracized to being seriously injured. As such, it is essential to understand how social and intergroup contexts bring about variations in language behaviors.

Studying how social context influences language behaviors between individuals in contact with each other is one of the essential goals of linguistic communication research. For example, Oskaar (1996) describes language contact phenomena as being greatly influenced by the sociocultural context in which speakers interact. By placing social contexts as forefront players in communication behavior, Oskaar reminds communication researchers that language behaviors cannot be understood in isolation, but should instead be analyzed within their broader social contexts.

One specific discipline that offers insight as to how social contexts impact individuals' contact behaviors is social psychology. While acknowledging the power of internal states in producing behaviors and attitudes, social psychologists emphasize how a person's attitudes and behaviors are greatly molded by contextual factors (Myers and Spencer 2006), particularly intergroup contexts. Because of its emphasis on social and intergroup dynamics, social psychology is well positioned to study how social context will impact the language used between individuals in contact. Specifically, the social psychology of language offers clear intergroup-based theories and predictions that can be used to understand why, when, where, and how individuals choose to communicate, as well as the consequences of these communications.

The present chapter presents an overview of the research on the social psychology of language that has received the strongest empirical support. First, we describe the importance of language as a symbol of one's social identities (Shepard, Giles, and Le Poire 2001). Understanding the symbolic value of language for groups is the basis for under-

standing language variation(s) as a function of the intergroup context. Second, we delve into how intergroup contexts can produce fluctuations in language behavior when communicating with another person (Giles et al. 1987) and when describing other individuals (Maass 1999). Finally, we examine how status is an important sociostructural intergroup variable that impacts how we learn new languages (Taylor and Wright 2002).

2. Language as a symbol of social identity

Languages, accents, dialects, or even specific words can be important markers of group identity. For example, in Quebec, Canada, the French language is used to mark the distinction between Québécois culture and mainstream Anglophone-Canadian culture (Bourhis 2001). Indeed, although not an exclusive, or even essential, marker of group identity, language behaviors are nonetheless often presented by cultural groups (Shepard, Giles, and Le Poire 2001), gender groups, age groups, and professional groups (Giles, Reid, and Harwood 2010) as a symbol of special significance. These distinct language behaviors become characteristics of a group's identity, and help distinguish its members from outsiders. For instance, when men interact with women, they adopt more power-related language behavior than when they interact with other men (Mulac 1989). Clearly, there are strong ties between language behavior and group identity.

Yet the question remains why language behavior is often seen as an important marker of group identity. Research suggests two reasons. First, language behavior is a distinct group feature that is shared by most in-group members, but not necessarily shared with members of other groups (Giles, Coupland, and Coupland 1991). According to social identity theory (Tajfel 1974), characteristics that are seen in most members of a given group, but not necessarily in other related groups, can easily come to characterize the group itself. As such, the language behaviors that we use become a source of information about one's group membership, signaling to others that we belong to a certain group. Furthermore, group characteristics, such as language, are usually seen by members of the group as being inherently positive, making them even more central to the group's identity. For example, Scottish teenagers from Dundee rated the Dundee accent as more favorable than the Scottish accent from another city and the British English accent (Abrams and Hogg 1987), even if objectively, no one accent is 'better' than another.

Secondly, language is an important marker of social identity because identities are shaped through conversations, that is, by having discussions about the meaning of such identities with fellow group members (Hecht, Jackson II, and Pitts 2005; Verkuyten 2010). Because language behaviors are necessary to discuss the attributes of a group's identity, they can become inherently tied to the group's identity itself. The importance of language when creating a socially shared identity is well illustrated by the French-Canadian youth living in English-majority provinces (Dallaire 2003). These individuals are proficient in English and French, and use both languages when negotiating together the meaning of their Franco-Canadian identities. Because they use both languages when engaging in conversations about their linguistic identities, they can combine their identities in new ways, allowing for the emergence of a new hybrid identity (as opposed to a single Francophone identity).

Since language behaviors are intimately tied to social identity, individuals who are conversing are necessarily giving each other cues about their group memberships and

social identities (Shepard, Giles, and Le Poire 2001). In other words, they communicate through their language behaviors whether or not they belong to the same social group. That is not to say that every conversation will necessarily be imbued with intergroup and social identity issues; however, every interaction can potentially become a conversation between representatives of the same or of different groups. Simply put, any conversation can be impacted by intergroup dynamics. Even something as mundane as a routine traffic stop by a policeman can be seen as an inherently intergroup interaction (civilian versus policeman; Giles, Choi, and Dixon 2010). In what follows, we discuss three language behaviors and theories that highlight how our social identities and intergroup contexts impact how we communicate with others, how we describe others, and how we learn languages.

3. Language and intergroup context: a social psychological approach

3.1. Intergroup contexts and how we communicate with others: convergence and divergence

Individuals adjust their language behaviors in reaction to the other person in the conversation (Shepard, Giles, and Le Poire 2001). When two individuals are conversing, such as an Anglophone and a Francophone in Montreal speaking in French about the weather, the Francophone could use some English words, speak slower, and even mimic their conversation partner's posture. These are examples of *convergence*, or the act of adjusting one's communicative behaviors to be more similar to the other (Dragojevic, Gasiorek, and Giles 2016). In contrast, individuals conversing together could *diverge* by using communicative behaviors that are dissimilar to their conversation partner (Dragojevic, Gasiorek, and Giles 2016), usually in order to exaggerate the level of distance between them. A diverging Francophone could use complex meteorological terminology in French, increase the speed of their speech, and thicken their French accent. Convergence and divergence are the most well-studied *accommodation behaviors* (Dragojevic, Gasiorek, and Giles 2016), but other accommodation behaviors have been identified (e.g. speech complementarity and maintenance).

Several theories have been proposed to explain why individuals engage in accommodation behaviors, particularly from an interpersonal perspective. For example, the discrepancy-arousal theory (Capella and Greene 1984) emphasizes the role of personal expectations (i.e. what one expects the other person in the conversation to do) when deciding which language behaviors to use. However, most theories focus exclusively on interpersonal relations, disregarding the relationship between language behavior and social identity, and thereby ignoring how each conversation can be potentially impacted by intergroup dynamics (Shepard, Giles, and Le Poire 2001). Communication accommodation theory (or CAT) emerged to explain accommodation behaviors by simultaneously considering interpersonal and intergroup relations, contexts, and motivations.

The basic tenet of CAT is that an individual's personal and group motivations will determine whether they will converge or diverge (Shepard, Giles, and Le Poire 2001; Thakerar, Giles, and Cheshire 1982). According to CAT, three motivations promote con-

vergence. First, convergence is expected when individuals are motivated to gain the approval of their conversation partner. By employing behaviors that are similar to their conversation partner, individuals promote 'liking' (Putman and Street 1984; see also Giles, Taylor, and Bourhis 1973). Second, individuals converge to indicate a common social identity between speakers. A common identity fosters liking and a sense of shared community, which can have important consequences (e.g. ordering coffee with the right accent; Dragojevic, Gasiorek, and Giles 2016). Third, individuals will converge when they want to facilitate communication with each other, e.g. when they want to communicate directions to the closest restroom.

In terms of divergent accommodation behaviors, CAT postulates that divergence occurs when individuals wish to emphasize their distinct personal and/or group identity. When seeking to reinforce their personal identity, individuals will use the speech characteristics that make them different from their conversation partner. When seeking to reinforce their group identity, individuals will engage in language behaviors which are distinctive of their own group and mark it as different from the conversation partner's group. More specifically, since language behaviors are important, valued markers of group identity, individuals can use the language characteristics of their group to mark their distinct and positive group identity in intergroup contexts (social identity theory; Tajfel 1974). Hence, an individual will diverge when seeking to clearly establish their distinct group membership. In a classic study on divergent behavior, Bourhis and Giles (1977) found that Welsh individuals adopted a stronger Welsh accent when they were confronted (i.e. when in an intergroup context) with an English man who was particularly negative towards the Welsh language.

In sum, CAT predicts that personal and social motivations dictate how individuals will communicate with each other, that is, whether individuals will want to converge or diverge. By combining both interpersonal and intergroup factors, CAT can pinpoint how our personal identities, social identities, and social contexts all color our conversations. However, the CAT framework is less useful when attempting to understand another communicative phenomenon: the way in which individuals talk about or describe each other. In the following section, we explore how verbal descriptions of others are impacted by intergroup contexts.

3.2. Intergroup contexts and how we describe each other: linguistic expectation bias and linguistic intergroup bias

An observation of, for example, an event, a behavior, and/or a person can be described in many different ways. For instance, a person running in a park might be described by an observer as a suspicious character disturbing a family environment, or as a health-conscious individual doing cardiovascular exercise. Based on the 'objective' scenario that was presented (a person running in a park), either description could be correct, and yet the outcome of each description is very different. If the observer describes the runner as a suspicious character, they are more likely to contact the police. If the runner is described as a health-conscious individual, the observer might not take any particular action. There are important cognitive, emotional, and behavioral consequences of how we describe each other.

One way in which the observer generates descriptors of the runner is by referencing the expected attributes of the group(s) to which the runner belongs, or in other words, by making use of group stereotypes. Indeed, assigning a person to a group (i.e. categorizing the person as a runner) is enough to generate the stereotypes associated with that group (i.e. the set of expected behavior[s] that should be observed in runners as a group). These stereotypes are used to understand the actions of the individual, even without specific knowledge about the individual themselves. The field of social psychology has long studied the origins and consequences of stereotypes (for a review of research on stereotypes and prejudice, see Dovidio, Glick, and Rudman 2005), and has found that stereotypes can be transmitted in subtle and implicit ways, such as by the words chosen to describe a group.

According to the Linguistic Category Model (Semin and Fiedler 1988), descriptions can be placed on a continuum ranging from very concrete descriptions to very abstract descriptions. At the concrete end of the continuum are *descriptive action verbs* (e.g. the person *runs*), which refer to a specific behavior that is easily verifiable, and hence noninterpretive. *Interpretive action verbs* follow descriptive action verbs on the continuum, referring to broader categories of actions instead of the specific action observed (e.g. the person *exercises*). *State verbs* follow interpretive action verbs on the continuum and are more abstract, as they describe the state of the individuals observed (e.g. the person *cares for his health*), not simply their behavior. At the abstract end of the continuum are adjectives and adverbs (e.g. the person *is active*), which are not tied to any specific observable behavior and are therefore difficult to verify (i.e. do active people run?). Descriptions close to the concrete end of the continuum tend to represent states that are situationally bound and caused by external factors, whereas the motivation and personality of individuals cannot be inferred from their actions (i.e. we do not know why the person runs). When abstract descriptions are used, the action is more likely to be attributed to internal, enduring individual states (i.e. an active person is seen as someone who possesses an internal quality that motivates them to run).

When do individuals use more abstract versus concrete descriptions? The answer to this question was proposed by the linguistic expectancy bias theory (Wigboldus, Semin, and Spears 2000; Wigboldus, Spears, and Semin 2005). Linguistic expectancy theory postulates that observations which fit expectations (i.e. stereotypes in intergroup contexts) are more likely to be described using abstract language, as this level of abstraction accurately describes the enduring nature of the observation (Wigboldus, Semin, and Spears 2000; Wigboldus, Spears, and Semin 2005). For example, if the person running in the park has a normal weight and is dressed in running attire, this observation resembles the stereotype (or expectations) of runners, and hence will be described with adjectives or adverbs (e.g. the person *is active*). On the other hand, observations that do not fit the stereotype will be described at a concrete level, closer to the action verbs end of the continuum. If the runner has an unhealthy weight and is not dressed in exercise clothes, the fact that they are running will not fit the stereotype of runners, and hence is more likely to be described in concrete terms (e.g. the person *is running*).

Linguistic intergroup bias (Maass 1999; Maass, Ceccarelli, and Rudin 1996; Maass et al. 1989) takes this theory a step further by postulating that individuals' motivation for a positive and distinct identity (Tajfel 1974) prompts them to describe the actions of the in-group and the out-group so as to enhance their identity. This translates into four hypotheses, two concerning in-group descriptions and two regarding out-group descrip-

tions. First, when describing the in-group's positive actions, abstract words will be used. This abstract description conveys that one's group engages in positive actions because it is an inherently positive group. Second, the negative actions of the in-group are described with concrete words, suggesting that they are situationally bound. Third, the positive actions of the out-group are described with concrete words that emphasize how the circumstances are responsible for their positive actions. Fourth, abstract words will be used to describe the negative actions of the out-group, as these reflect the enduring and internal negative qualities of this group. Results from various studies support these hypotheses, particularly those pertaining to descriptions of the positive actions on the part of the in-group versus the out-group (Maass, Ceccarelli, and Rudin 1996; Maass et al. 1989). Importantly, research shows that such linguistic biases were absent when individuals were in interpersonal contexts, highlighting their unique usefulness in intergroup contexts (Wigboldus, Spears, and Semin 2005).

The stereotypes we hold and verbalize in a conversation also carry important information concerning our own group identity. An individual who uses abstract descriptions of their group's positive actions and concrete descriptions of its negative actions was more readily signaled by observers as a member of that group (Porter, Rheinschmidt-Same, and Richeson 2016), and was also judged as a better member of said group (Assilaméhou and Testé 2012). As such, the way we describe others highlights our own group identity.

The implications of linguistic intergroup bias and linguistic expectancy bias for communication researchers are clear. This research highlights how one's group memberships lead to bias in one's descriptions (see also Sutton 2010). Second, they highlight how descriptors help create and perpetuate stereotypes and prejudice (Maass et al. 1989). Indeed, an individual who describes an out-group with negative abstract behavior is not only signaling their negative view of the out-group, but also predisposing the hearer to seeing the out-group as holding these underlying negative attributes (Karpinsky and von Hipper 1996; see also Wigboldus and Douglas 2007). As such, language is a path for perpetuating stereotypes while maintaining group and language status. In the following section, we explore how language status can impact one of the most fundamental elements of communication: how we learn a language.

3.3. Intergroup contexts and how we learn languages: language status

Bilinguals and multilinguals form a majority in the world. Today, it is estimated that one half to two thirds of the world's population knows two or more languages (Baker 2011: 66). As globalization continues, rates of bilingualism will only increase. This means that individuals are more and more likely to learn the languages − and hence the key markers of social identity − of other groups. As this trend continues to grow, it becomes important to understand the consequences of learning languages other than one's own mother tongue. Key findings in bilingual research indicate that individuals may experience either additive bilingualism or subtractive bilingualism (Baker 2011; Lambert 1975). *Additive bilingualism* occurs when the mother tongue abilities of the learner are not impacted by the acquisition of the new language. Hence, the individual is adding a new language without any consequence on their mother tongue. In contrast, s*ubtractive bilingualism* takes place when mother tongue competency decreases while individuals become in-

creasingly proficient in the newly acquired language. By adding a new language, the individual is subtracting from their mother tongue's abilities.

Some researchers argue that the occurrence of additive versus subtractive bilingualism depends on the status attributed to the learner's group of origin (Wright and Taylor 1995). Group status results from the human tendency to organize groups and individuals into a hierarchy (Pratto 1999). At the top of the hierarchy are those who have valued characteristics (e.g. education, richness, and power), and thus are conferred high status, while those groups who have little of the valued characteristics are low in status. Group status affects intergroup dynamics because status is positively linked to competency, privilege, and accessibility to resources, all characteristics sought by individuals and groups (Fiske 2010). Hence, the status of the language being learned compared to the status of the native language may lead to different patterns of bilingualism, as the motivations to learn the language will be different.

Individuals that are part of a high-status group will seek to uphold their group's position in order to maintain the advantages that stem from their membership (Chang 2015). Thus, when an individual from a high-status group learns a new language, they will be motivated to keep their mother tongue because it is a marker of their high-status group membership. Hence, their mother tongue abilities will be preserved while learning the new language. An example of additive bilingualism is seen in French immersion programs in Canada (Swain and Lapkin 1991). English is the language spoken by the majority in Canada, featured prominently in the arts, media, and academia; as such, English is generally considered to have higher status than French. In this context, Anglophone students in French immersion programs maintain competency in English while becoming more fluent in French, exhibiting additive bilingualism (Swain and Lapkin 1991).

In contrast to high-status group members, members of low-status groups are more open to adopting the behaviors of the high-status group to increase their own status (Tajfel and Turner 1986). As a marker of social identity, learning the language of a higher status group can increase one's social standing, and lead to a preference for the high-status language over one's own low-status mother tongue. As such, for individuals of low-status groups, subtractive bilingualism can result when learning the languages of higher status groups. Research with different minority groups of low status, such as the Inuit in Québec (Wright, Taylor, and Macarthur 2000) and Franco-Americans in Maine (Landry and Allard 1992), has found that low-status minorities show subtractive bilingualism when educated in the language of a high-status majority.

A deeper understanding of how status leads to additive versus subtractive bilingualism offers insights into how the larger social context and intergroup relations influence bilingualism (Landry and Allard 1992). Members of higher status groups are often favored in their bilingualism by gaining an additional language; individuals of lower status groups risk losing their mother tongue abilities when becoming bilingual. Knowing the impact of status on language, appropriate actions may be taken to favor additive processes among low-status members. For instance, educational programs that give equal status to both the minority and the majority languages within the school have been found to promote additive bilingualism in minority members (e.g. Murphy 2014). By understanding the power of status, linguistic minorities may thus be able to protect their language (Lambert 1975).

4. Conclusion

When two or more individuals are in contact, each of them brings to the encounter a rich history forged by their shared and respective social contexts, i.e. they bring their social identities along with their personal characteristics. One's social identity becomes increasingly salient when in contact with others, as individuals compare how similar to, or different from, each other they are. A conversation between two individuals may quickly become a conversation between representatives of two genders, occupations, sports teams, or ethnicities. Therefore, to better understand how language contact influences the way we communicate with each other (i.e. the language behaviors we display), it is imperative to consider how the social identity of individuals comes into play.

Ultimately, the goal of understanding language and contact through the scope of social psychology is to understand how intergroup relations and the social identity of individuals impact mutual communication and understanding. In this chapter, we highlighted three communication behaviors and forms of contact that are influenced by social identity: accommodation behavior, descriptions of observations, and language acquisition. These behaviors, and the social psychology theories developed to investigate them, shed light on how social context and intergroup relations influence the way individuals in contact with each other communicate.

Research on language behavior is increasingly taking a social identity perspective. For example, research on labeling shows that our attitudes towards a migrant depend on whether the person is labeled using a single label (e.g. French) or dual (or hyphenated) label (e.g. French-Canadian [Verkuyten and Thijs 2010]). Since the dual labeled groups normally carry the name of the majority group, they are more likely to be well received by the majority, which can promote positive intergroup relations. More research on language contact that considers the central role of the social identity of individuals could therefore promote favorable intergroup relationships and explain situations where such relationships are not the case.

Our hope is that this broad overview of the social psychology of language will relay to the reader how central intergroup contexts are in forging positive versus negative contact experiences. As such, it becomes pivotal to consider the social identity of individuals when studying such behaviors, especially considering how globalization increasingly encourages contact between different linguistic groups.

5. References

Abrams, Dominic & Michael A. Hogg
1987 Language attitudes, frames of reference, and social identity: A Scottish dimension. *Journal of Language and Social Psychology* 6(3–4). 201–213.

Assilaméhou, Yvette & Benoit Testé
2012 The effects of linguistic abstraction on evaluations of the speaker in an intergroup context: Using the linguistic intergroup bias makes you a good group member. *Journal of Experimental Social Psychology* 49(1). 113–119.

Baker, Colin
2011 *Foundations of bilingual education and bilingualism.* Clevedon: Multilingual Matters.

Bourhis, Richard Y.
2001 Reversing language shifts in Quebec. In Joshua A. Fishman (ed.), *Can threatened languages be saved? Reversing language shift revisited: A 21st century perspective*, 101–141. Clevedon: Multilingual Matters.
Bourhis, Richard Y. & Howard Giles
1977 The language of intergroup distinctiveness. In Howard Giles (ed.), *Language, ethnicity, and intergroup relations*, 119–134. London: Academy Press.
Capella, Joseph N. & John O. Greene
1984 The effects of distance and individual differences in arousability on nonverbal involvement: A test of discrepancy-arousal theory. *Journal of Nonverbal Behavior* 8(4). 259–285.
Chang, Jin W.
2015 *The effects of group status on intragroup behavior: Implications for group process and outcome*. Pittsburgh, PA: Carnegie Mellon University MA thesis.
Chomsky, Noam
1957 *Syntactic structures*. The Hague: Mouton.
Dallaire, Christine
2003 "Not just francophone": The hybridity of minority francophone youths in Canada. *International Journal of Canadian Studies* 28. 163–199.
Dovidio, John F., Peter Glick & Laurie A. Rudman
2005 *On the nature of prejudice: Fifty years after Allport*. Malden: Blackwell.
Dragojevic, Marko, Jessica Gasiorek & Howard Giles
2016 Communication accommodation theory. In Charles R. Berger & Michael E. Roloff (eds.), *The international encyclopedia of intergroup communications*, 1–21. Hoboken: Wiley-Blackwell.
Fiske, Susan T.
2010 Interpersonal stratification: Status, power, and subordination. In Susan T. Fiske, Daniel T. Gilbert & Gardner Lindzey (eds.), *Handbook of social psychology*, 5th edn., 941–982. Hoboken: Wiley.
Giles, Howard, Anthony Mulac, James J. Bradac & Patricia Johnson
1987 Speech accommodation theory: The first decade and beyond. *Annals of the International Communication Association* 10(1). 13–48.
Giles, Howard, Charles W. Choi & Travis L. Dixon
2010 Police-civilian encounters. In Howard Giles, Scott Reid & Jake Harwood (eds.), *The dynamics of intergroup communication*, 65–76. New York: Peter Lang.
Giles, Howard, Donald M. Taylor & Richard Y. Bourhis
1973 Towards a theory of interpersonal accommodation through language: Some Canadian data. *Language in Society* 2(2). 177–192.
Giles, Howard, Nikolas Coupland & Justine Coupland
1991 Accomodation theory: Communication, context, and consequence. In Howard Giles, Justine Coupland & Nikolas Coupland (eds.), *Contexts of accommodation: Developments in applied sociolinguistics*, 1–68. Cambridge: Cambridge University Press.
Giles, Howard, Scott Reid & James Harwood (eds.)
2010 *The dynamics of intergroup communication*. New York: Peter Lang.
Hecht, Michael L., Ronald L. Jackson II & Margaret J. Pitts
2005 Culture: Intersections of intergroup and identity theories. In Jake Harwood & Howard Giles (eds.), *Intergroup communications: Multiple perspectives*, 21–42. New York: Peter Lang.
Karpinsky, Andrew & William von Hipper
1996 The role of the linguistic intergroup bias in expectancy maintenance. *Social Cognition* 14(2). 141–163.

Lambert, Wallace E.
1975 Culture and language as factors in learning and education. In Aaron Wolfgang (ed.), *Education of immigrant students*, 55–83. Toronto: Ontario Institute for Studies in Education.

Landry, Rodrigue & Réal Allard
1992 Subtractive bilingualism: The case of Franco-Americans in Maine's St John Valley. *Journal of Multilingual & Multicultural Development* 13(6). 515–544.

Maass, Anne
1999 Linguistic intergroup bias: Stereotype-perpetuation through language. *Advances in Experimental Social Psychology* 31. 79–121.

Maass, Anne, Daniela Salvi, Luciano Arcuri & Gün Semin
1989 Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology* 57(6). 981–993.

Maass, Anne, Roberta Ceccarelli & Samantha Rudin
1996 The linguistic intergroup bias: Evidence for ingroup-protective motivation. *Journal of Personality and Social Psychology* 71(3). 512–526.

Mulac, Anthony
1989 Men's and women's talk in same-gender and mixed-gender dyads: Power or polemic? *Journal of Language and Social Psychology* 8(3–4). 249–270.

Murphy, Audrey F.
2014 The effect of dual-language and transitional-bilingual education instructional models on Spanish proficiency for English language learners. *Bilingual Research Journal* 37(2). 182–194.

Myers, David M. & Steven J. Spencer
2006 *Social psychology*, 3rd edn. New York: McGraw-Hill.

Oskaar, Els
1996 Conceptual prerequisites of contact linguistics I: Central issues of the history of contact linguistics. In Hans Goebl, Peter H. Nelde, Zdenek Stary & Wolfgang Wölck (eds.), *Contact linguistics: An international handbook of contemporary research*, vol. 1, 1–12. Berlin: Walter de Gruyter & Co.

Porter, Shanette C., Michelle Rheinschmidt-Same & Jennifer A. Richeson
2016 Inferring identity from language: Linguistic intergroup bias informs social categorization. *Psychological Science* 27(1). 94–102.

Pratto, Felicia
1999 The puzzle of continuing group inequality: Piecing together psychological, social, and cultural forces in social dominance theory. *Advances in Experimental Social Psychology* 31. 191–263.

Putman, William B. & Richard L. Street Jr.
1984 The conception and perception of noncontent speech performance: Implications for speech-accommodation theory. *International Journal of the Sociology of Language* 46. 97–114.

Robinson, W. Peter
2001 Language in communication: Frames of reference. In W. Peter Robinson & Howard Giles (eds.), *The new handbook of language and social psychology*, 3–32. West Sussex: Wiley.

Semin, Gün R. & Klaus Fiedler
1988 The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology* 54(4). 558–568.

Shepard, Carolyn A., Howard Giles & Beth A. Le Poire
2001 Communication accommodation behavior. In W. Peter Robinson & Howard Giles (eds.), *The new handbook of language and social psychology*, 33–56. West Sussex: Wiley.

Sutton, Robbie M.
2010 The creative power of language in social cognition and intergroup relations. In Howard Giles, Scott Reid & Jake Harwood (eds.), *Dynamics of intergroup communication*, 105–115. New York: Peter Lang.
Swain, Merril & Sharon Lapkin
1991 Additive bilingualism and French immersion education: The roles of language proficiency and literacy. In Allan G. Reynolds (ed.), *Bilingualism, multiculturalism, and second language learning*, 203–216. Hillsdale: Lawrence Erlbaum.
Tajfel, Henry
1974 Social identity and intergroup relations. *Social Sciences Information* 13(2). 65–93.
Tajfel, Henry & John C. Turner
1986 The social identity theory of intergroup behavior. In Stephen Worchel & William G. Austin (eds.), *The psychology of intergroup relations*, 7–24. Chicago: Nelson-Hall.
Taylor, Donald M. & Stephen C. Wright
2002 Do aboriginal students benefit from education in their heritage language? Results from a ten-year program of research in Nunavik. *Canadian Journal of Native Studies* 22(1). 141–164.
Thakerar, Jitendra N., Howard Giles & Jenny Cheshire
1982 Psychological and linguistic parameters of speech accommodation theory. In Colin Fraser & Klaus R. Scherer (eds.), *Advances in the social psychology of language*, 205–255. Cambridge: Cambridge University Press.
Verkuyten, Maykel
2010 Ethnic communication and identity performance. In Howard Giles, Scott Reid & James Harwood (eds.), *The dynamics of intergroup communication*, 17–28. New York: Peter Lang.
Verkuyten, Maykel & Jochem Thijs
2010 Ethnic minority labeling, multiculturalism, and the attitude of majority group members. *Journal of Language and Social Psychology* 29(4). 467–477.
Wigboldus, Daniël, H. J. & Karen M. Douglas
2007 Language, stereotypes, and intergroup relations. In Klaus Fiedler (ed.), *Social communication*, 79–106. New York: Psychology Press.
Wigboldus, Daniël H. J., Gün R. Semin & Russell Spears
2000 How do we communicate stereotypes? Linguistic basis and inferential consequences. *Journal of Personality and Social Psychology* 78(1). 5–18.
Wigboldus, Daniël H. J., Russell Spears & Gün R. Semin
2005 When do we communicate stereotypes? Influence of the social context on the linguistic expectancy bias. *Group Processes and Intergroup Relations* 8(3). 215–230.
Wright, Steve C. & Donald M. Taylor
1995 Identity and the language of the classroom: Investigating the impact of heritage versus second language instruction on personal and collective self-esteem. *Journal of Educational Psychology* 87(2). 241–252.
Wright, Steve C., Donald M. Taylor & Judy Macarthur
2000 Subtractive bilingualism and the survival of the Inuit language: Heritage- versus second-language education. *Journal of Educational Psychology* 92(1). 63–84.

Diana Cárdenas, Utrecht (The Netherlands)
Laura French Bourgeois, Montréal (Canada)
Roxane de la Sablonnière, Montréal (Canada)

64. Translation studies

1. Common ground (in principle)
2. A striking lack of 'contact' (in practice), or at least until recently
3. Language contact in TS: conceptualizations and perspectives
4. References

1. Common ground (in principle)

In the past half century or so, Translation Studies (TS), which will be used here to encompass Interpreting Studies (IS) as well (unless otherwise stated), has progressively established itself as an autonomous discipline, albeit one with relatively vague contours and an outspoken interdisciplinary or transdisciplinary status. Its vitality is demonstrated by a rapidly increasing research output, much of which sees the light of day in the many dozens of journals or several book series created mostly in the past two decades. Guidance through this mass of scholarly writing can be obtained from bibliographies such as BITRA (open access, hosted by the University of Alicante) and the Translation Studies Bibliography (by subscription, published by John Benjamins). Furthermore, students, teachers, and researchers can benefit from a growing number of handbooks (e.g. Bermann and Porter 2014; Gambier and van Doorslaer 2010–2013; Malmkjaer and Windle 2011) and textbooks (e.g. Munday 2012; Pym 2014). Venuti (2012) is an influential anthology of key papers.

The central research object of TS, i.e. translation processes and the products resulting from them, can hardly be described otherwise than as a particularly intense form of interlingual contact. For that reason alone, TS would appear to prominently stand out as a privileged partner discipline for Contact Linguistics (CL). Not surprisingly, some of the recurrent research themes in TS show strong overlap with CL research priorities. Let us quote two examples of this.

First, TS has shown since its beginnings an awareness of the obvious fact that translation is only one of several possible strategies to which speakers have recourse when different languages come into contact – and that it should be conceptualized as such. The other options include foreign language learning, code-switching, use of a *lingua franca*, etc., all of which belong to the core topics of CL. Already in the early 1980s, Koller (1983: 20–27) duly situated translation among a range of other "*Möglichkeiten der Überwindung von Sprachbarrieren*" ('options for overcoming language barriers'). Holmes ([1972] 1988: 77–78) had previously reserved a niche for policy-related matters in the 'applied' branch of his blueprint of what the new discipline of TS would have to look like:

> [a] third area of applied translation studies is that of *translation policy*. The task of the translation scholar in this area is to render informed advice to others in defining the place and role of translators, translating, and translations in society at large: […] determining what works need to be translated in a given socio-cultural situation, what the social and economic position of the translator is and should be, or […] what part translating should play in the teaching and learning of foreign languages.

https://doi.org/10.1515/9783110435351-064

It is significant that Holmes had the foresight to include in his foundational 'map' of the discipline the task of defining translation policies on the basis of societal, economic, and didactic considerations. It was left to later scholars to discover how translation policies and other language policies are inextricably interwoven, but TS scholars are taking up the gauntlet (e.g. González Núñez 2016; Pym 2008), showing that the study of translation is not confined to what happens once translation is underway, but also includes the very conditions and prior motivations that led to translational action being taken (or *not* being taken) in a given language contact situation.

A second example would be the phenomenon of linguistic interference, which has for more than half a century (starting at least with Weinreich's work) belonged to the conceptual toolbox of CL, but which is no less a central concern in TS, especially in its more linguistically (and often didactically and prescriptively) inspired forms. Work in TS from the 1980s–1990s by Koller, Albrecht, or Neubert might be quoted here; see also Schmidt (1989). There is indeed no denying that various 'shining-through' phenomena tend to pervade the performance of translators in a way which directly recalls interference in second language acquisition, contact-induced linguistic change, and several other phenomena at the heart of CL's research agenda. Interestingly, interference-like phenomena not only appear in the 'typical' case of the amateur or novice translator, but may also reflect a more deliberate (e.g. artistically, ethically, and/or politically motivated) commitment to achieve 'foreignization' by breaking the illusion of fluency and transparency, thus highlighting the intervention of the translator, as well as the 'otherness' of the source text, its language, and its culture. Here, as with the first example above, one notes a general shift from broadly normative attitudes, with interference initially merely frowned upon as a 'negative' form of transfer and a source of 'inaccurate' and 'unidiomatic' translations, to more descriptive (or at least, more context-sensitive) approaches.

2. A striking lack of 'contact' (in practice), or at least until recently

Nevertheless, and strangely enough, one finds hardly any mention of CL in the mainstream TS literature. None of the five handbooks and textbooks listed in our first paragraph above mentions CL as such in its index. Admittedly, *The handbook of translation studies* (Gambier and van Doorslaer 2010–2013) has an entry on "Translation zone" (Simon 2013), which duly refers to Pratt's (1991) work on 'contact zones'; the same handbook also features an entry on "Multilingualism and translation" (Meylaerts 2010). Taken together, however, these reference works illustrate the worrying extent to which TS has been oblivious of the very existence of CL as a sister discipline.

Conversely, CL has traditionally also shown little interest in the achievements of TS. For example, the survey chapter prefacing Winford's *An introduction to contact linguistics* (2003: 1–28) does not so much as mention the word 'translation'. As Kranich (2014: 97) has put it more recently, "[t]he major studies of language contact only discuss language contact through translation (LCTT) very briefly […] or even merely mention it in passing". Kranich (2014) is itself a fine effort to promote the dialogue between TS and CL, but its scope remains limited to the diachronic study of translation-induced linguistic change and variation, which covers only part of the potential interface between CL and TS. In 2014 the Call for Papers for a conference on "Language contact: *The state of*

the art [emphasis mine]" to be held in Helsinki mentioned the possible theme "current developments in the study of the so far *rather under-explored* [emphasis mine] modes of language contact, e.g. translating, interpreting" (see www.linguistics.fi/contact/ − last access 15 March 2017). We may trust that the phrase "rather under-explored" bears upon translation-related research *done within CL*, and does not reveal the organizers' ignorance of work done *in TS*, but even so, the phrase illustrates the startling scale of the mutual neglect between CL and TS.

The situation just described raises further questions about the exclusionary effects that follow from the zeal of communities of researchers working in a given field to foster their own 'discipline' by defining its own conceptual and methodological paradigms and creating separate publication outlets, and about the challenges awaiting any cross-disciplinary endeavors to counter such exclusionary effects. A recent collection of papers entitled *Border crossings: Translation studies and other disciplines* (Gambier and van Doorslaer 2016) aims to address such issues as they apply to TS − a field that has been very much beset by them, having been described as a somewhat disjointed 'polydiscipline' in its earlier stages, before it became a more integrated 'interdiscipline', and more recently, perhaps even a 'transdiscipline'. The volume in question does not explicitly mention CL, but it does contain a chapter entitled "Multilingualism studies and translation studies: Still a long road ahead" (Meylaerts and du Plessis 2016). While this title implies a downbeat appraisal of the current state of affairs, the chapter opens interesting avenues, and is in itself a sign of hope, provided we all agree not to become lost in territorial debates about the borderlines between the overlapping fields of CL and multilingualism studies, to name but those. Interdisciplinarity has been a trigger for change here, and it is very encouraging that its spirit also pervades a recent volume on the CL side of the divide, significantly entitled *Language contacts at the crossroads of disciplines* (Paulasto et al. 2014) and containing a section on translation.

Looking back, one is tempted to assume that the near-absence (until recently) of a broad-based, direct dialogue between TS and CL is a matter of a well-reasoned division of labor. Various distinctions might be invoked to support such a belief. TS could thus be said to focus more on texts (versus language users in CL), written language (versus spoken language), language professionals (versus general language users), and/or cultural aspects (versus more strictly linguistic ones), yet none of these discriminations really hold water in the final analysis. The sobering fact remains that the prevailing mutual ignorance is largely an affair of counterproductive academic compartmentalization, itself the result of a complex interplay of intellectual, sociological, psychological, institutional, and increasingly, economic forces. A parallel and comparative history of the two fields could shed more light on this matter, but is beyond the scope of this chapter (however, see D'hulst and Gambier 2018).

Turning our eyes to the future, the main thing that counts is that both TS and CL now appear to have their feelers out, and seem ready for mutual exploration and interaction. To the signs of hope on the side of TS that we have already mentioned, we could further add, with more examples to follow, a recent article on "Understanding translation as a site of language contact" (Malamatidou 2016), as well as Angelelli's (2016) chapter on "Bilingualism and multilingualism" in a methodological handbook addressed to a readership of TS students and scholars (Angelelli and Baer 2016).

Things have started to move on the CL side, too, where we have seen recent research giving greater prominence to translation in studies on themes such as the lasting dia-

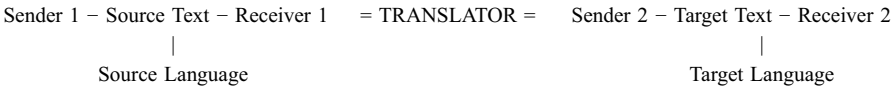
chronic effects of translations, synchronic linguistic variation caused by translation contacts, loan translation, and the nature of translational contact as compared to other types of language contact. Kranich et al. (2011), which brings together approaches by historical linguists, language contact researchers, and translation scholars, has been influential in outlining this emerging field of study in CL. Significantly (and encouragingly), several scholars (including Baumgarten, House and Kolehmainen, to name but a few) have published work in both CL and TS outlets.

3. Language contact in TS: conceptualizations and perspectives

This entry cannot aspire to give a full panoramic survey of topics and issues in TS. We shall offer a brief survey of the history and range of TS, and try to highlight a few striking features of TS research that will hopefully suggest possibilities for fruitful interactions with CL. Some areas, where overlap and the beginning of a rapprochement between CL and TS may be noted, have already been mentioned (and more will follow), but we shall also mention areas of apparent divergence between TS and CL, assuming that TS' different conceptualizations of language contact may inspire future work in CL, whether directly or indirectly.

3.1. Linguistic equivalence and beyond

While translation theory has a very long history, stretching back at least as far as Cicero, 'modern' and more 'scholarly' theories of translation started to emerge in the 1950s–1960s. Having grown, in part, out of Applied Linguistics and the search for fully automatic translation, while also reflecting the postwar ethos of worldwide understanding and universal communication, early models in TS were dominated by the concept of translational equivalence. This concept has been variously defined, but in earlier stages of TS, it was usually understood in strictly linguistic terms. The strength of the belief in the impartiality of translation contacts ('equivalence' as 'same value') may be illustrated by the standard visual representations of the translation process, which invariably show some variation on the following:



This neatly balanced diagram is essentially a reduplicated version of traditional monolingual models of communication. It locates the contact between the source and the target languages (and between the initial sender and the ultimate receiver) in the individual translator, whose interlinguistic skills should enable them to achieve equivalence in each particular case. This model conceptualizes translation as a highly regulated and restricted form of language contact, reducing it, on the basis of an adequate matching of the respective linguistic systems, to a decoding/recoding procedure and the transfer of a message. In order to perform this task expertly, translators must have advanced levels

of bilingualism (a theme obviously at the interface of TS and CL), but they also must develop an extra set of specific skills that 'ordinary' bilinguals do not normally have, providing the rationale for a specialized translation didactics.

Since the 1970s, TS scholars have emphasized more and more the need to frame such narrowly linguistic understandings within broader contexts, thereby giving greater attention to a range of functional considerations, such as genre-related and/or pragmatic ones. Such considerations also later came to acknowledge the importance of cultural and political factors. In the 1990s, this trend took the form of what is often referred to as the 'cultural turn' in TS (see Section 3.2. below). This went so far as to generate an interest in so-called 'cultural translation', a somewhat nebulous notion that may extend to just about any form of intercultural movement, communication, or understanding, whether or not translation in the strict linguistic sense is involved.

This general trend urges the researcher to understand translation as a contact phenomenon between cultures (identities, and/or ideologies), as well as between language systems. Major first steps in this direction were taken by a group of scholars in the later 1970s and 1980s, who took cues from polysystem theorists such as Even-Zohar and Toury, who had in turn been influenced by Russian formalism and the Prague School. It is worth noting here that Even-Zohar was fully conversant with Weinreich's work on language interference, and went on to develop models of cultural transfer and planning. His long-term scholarly project could arguably be construed as an effort to develop for literary and cultural studies what CL was trying to achieve for linguistics. Even-Zohar takes contact and interference as key concepts in his thinking about TS, literary history, comparative literature, and cultural studies − fields which he tries to integrate into his research, following in the footsteps of Tynjanov and Jakobson within the larger tradition of "dynamic functionalism" (Even-Zohar 1990: 53). In one of his papers, Even-Zohar (1990: 53−72) even ventures to formulate a hypothetical set of "laws of literary interference", which implies the need to distinguish between different types of contacts between literatures, as not all contacts will necessarily lead to interference.

Translation was thus reframed within a theory of culture, not understood in static terms as a heritage, a canon, or an otherwise fixed structure, but rather viewed as an open, dynamic network of systems (and subsystems) which show hierarchies and tensions between norms and models striving for dominance. Noncanonized discourses, such as popular music, advertising, journalism, and so on, came to be included in research programs along with canonized ones, and translations were systemically correlated with nontranslated discourses. Choices regarding *whether* to translate at all, *what* to translate, and *how* to translate it are taken at the pole of the receptor ('target') cultural system, and these choices reflect the latter's functional needs, be these linguistic, cultural, or political. In other words, the modes and modalities of translation contacts are no longer prescribed by the scholar on the basis of the source texts' features, but rather empirically observed in given historical situations (hence the name Descriptive Translation Studies). Special attention is given to the receptor culture, whose norms of translation will determine if a target-language text resulting from a transfer operation shows the 'right' kind and amount of equivalence to its source in order to serve as a 'valid' translational substitute of an original.

This perspective is consistently functional. A pioneering example was Toury's research on the impact of translation in the evolution and revival of the Hebrew language and culture, which resulted in synchronic variety and diachronic change. The field of

translation history is replete with examples of how translations can have a tangible effect on the evolution of linguistic and literary repertoires. With 'weak' (e.g. young, peripheral, and/or unstable) cultures, translation strategies are often source-oriented (or 'foreignizing'), enabling more prestigious source-text linguistic and literary models to bring about change in the target culture. With 'stronger' receptor cultures, however, translation norms are more often target-oriented (or 'domesticating'), which means that translations are made to fit into the mold of the established target-culture models, such that their function is more to consolidate existing repertoires rather than to initiate or reinforce innovation in the target culture.

3.2. Contact metaphors in TS: the bridge, the conflict zone, the gap

Systemic thinking about translation, and its function within and between cultures, has the notion of struggle and domination built in; this notion has cast doubt on traditional conceptualizations of translation as an interlingual and intercultural 'bridge' permitting a harmonious meeting of minds and meanings somewhere in the neutral middle of a level playing field. The traditional equivalence-based model of translation as a discreetly efficient bridge-like facilitator of contacts has thus come to be contested by a model which views translational contacts as sites of conflicting interests (and often unbridgeable differences), which correspondingly require more interventionist (and possibly more 'visible') forms of negotiation on the part of the translator.

The notion that something like a 'cultural turn' happened in TS in the 1990s is specifically linked to the growing awareness of the power imbalance often involved in translational encounters, whether such power differentials must be defined in terms of the majority/minority status of the languages involved, political centers/peripheries, cultural prestige, gender, (post)coloniality, or in other ways. Many TS scholars have thus been drawn to the study of translation in big cities, borderlands, or creole-speaking communities, or in the lives of migrants, refugees, asylum seekers, or exiled authors. In the present context, it is worth noting that the theme of the status of minority languages as source and/or target languages in translation has attracted a significant amount of research attention (Branchadell and West 2005), to the point of prompting the creation in 2009 of a journal specifically dedicated to it. *mTm: A Translation Journal*, as it is called, aims at "starting and promoting a discussion on the particularities of translation from major into minor languages and vice versa, as well as of translation between minor languages" (see www.mtmjournal.gr).

The cultural turn has caused the locus of translational contact to be increasingly conceptualized as a zone of tension, competition, and conflict, where different worldviews, ideological positions, and interests meet (and potentially clash). Mona Baker's *Translation and conflict* (2006) has been an influential work in underscoring this nexus between power, translation, and conflict. Baker's book also illustrates another trend, namely, the move beyond descriptivism towards a robustly activist stance, which means that the researchers *themselves* relinquish any claims to neutrality, and instead show ethical and political commitment on behalf of oppressed or minoritized groups − a phenomenon which is not unknown in CL.

Yet another metaphorical conceptualization of the contact zone in translation construes it more as a 'gap' or an 'absence', implying that there can be no real contact

between meanings and languages: what we perceive is merely a socially constructed *illusion* of translingual, contact-enabling bridges being available to us. Such a metaphorization of translational contact underlies Derrida-inspired approaches to language and translation, which have been very influential in TS and whose leitmotif is the instability of meaning, in turn leading to unavoidable semantic slippage and dislocation in translation. Derrida rejects the Western philosophical tradition, which allegedly construes meaning in terms of some prelingual presence that can be expressed by the stable, closed, and homogenous structures of language, and that can be correspondingly carried across in its sameness between languages. Instead Derrida emphasizes absence, difference, and heterogeneity as existing within and (thus unavoidably) between languages, which ultimately implies a rejection of interlingual borderlines and of any binary distinction between translated and nontranslated discourse. Deconstruction understands the undecidability that follows from this argument not as a warrant for an 'anything goes' approach, but rather as actually increasing the translator's ethical responsibility in what is essentially an impossible task (Davis 2001). This paradox had been captured earlier in the title of Walter Benjamin's ([1923] 2012) famous essay "*Die Aufgabe des Übersetzers*", in which *Aufgabe* has the double meaning of 'task', and of 'giving up, abandonment'.

3.3. The multilayered nature of translational contact

TS scholars have increasingly shown an awareness of the importance of the complex genealogies of individual texts that connect the 'original' to its 'translation'. A modern translation of Shakespeare's *Hamlet*, for example, inserts itself into a massive web of intertextual relationships which stretches across centuries and includes: the early 17th century Quarto and Folio texts, the long editorial and critical traditions that have made these early versions of Shakespeare's work accessible and readable to later generations, and the many translations and adaptations of the play that already exist, thus requiring each new translation to position itself *vis-à-vis* earlier rewritings through positive and negative translational intertextuality by either embracing or rejecting the available models and readings. With older texts, retranslation thus becomes a standard feature. With source texts being rendered from a 'remote' source language into a 'small' target language, forms of indirect translation occur equally often, with the relay language usually being a larger, more prestigious one.

Such complex textual trajectories are not just typical of the translation of venerable or exotic texts from the past or from other continents, but are also a staple of everyday contemporary communicative practices. Examples abound in the world of journalism, e.g. where 'local' news is routinely translated into the language of major news agencies before being brought into international circulation, and translated again into the language of its ultimate consumers. These processes involve different agents (see Section 3.5 below) and often more than two languages, and they go hand in hand with other rewriting techniques (be they editorial ones, such as summarizing or commenting, or more overtly politically motivated ones, such as censorship or framing) in ways which mostly remain invisible at the point of reception. Whether it is Shakespeare or a news report, the standard binary source/target model of translation is far too simplistic to accommodate the complex configurations of agents and the sequences of intertexts which appear on

the scholar's radar and give the translation contact a composite or multilayered quality (for a recent collection of papers on this, see Cordingley and Frigau Manning 2017).

3.4. The mediated nature of translational contact

From its beginnings, TS has had no option but to develop a keen awareness of the mediated nature of translational communication, by which we mean the ways in which it is shaped by the material (physical and technological) channel being used. Indeed, it is on the basis of very pragmatic considerations, such as pedagogical and vocational ones, that the field soon had to acknowledge the distinction between (written) translation and (oral) interpreting. The media of speech and writing, respectively, come with their own material and semiotic constraints, which underlie the specificities of written compared to oral translation, each requiring its own skillset, technologies, and didactic approaches.

TS (here excluding interpreting) focused on *written* translation, which was often reduced to *literary* translation and to the study of how literary classics from the past were translated (or 'mistranslated') by later generations. This is a useful reminder that TS was also, in part, originally an offshoot of Comparative Literature. The Bible, too, was often a preferred object of study in the pioneering years of TS: witness the work of Eugene Nida and his associates, whose contributions to the field were framed within the missionary objectives of the American Bible Society and later the United Bible Societies. This reveals yet another of TS' many roots.

The remit of Interpreting Studies (IS) is to focus on *oral* translation, which in reality covers a wide range of practices, including conference interpreting and interpreting in a diversity of community settings, with sign language interpreting presenting yet another set of parameters and challenges. IS could be said to have started in the 1950s, building on the work of a handful of earlier pioneers, but it is not until the late 1980s and early 1990s that a strong research community developed, along with a sense of the field's distinct disciplinary status (Pöchhacker 2016). The study of interpreting now has a somewhat ambiguous relationship with respect to TS, with some still regarding IS as a subfield of TS, while others view it as an analogous, but autonomous, field in its own right. The name Translation and Interpreting Studies (abbreviated as TIS) is sometimes used to diplomatically emphasize both the commonalities between TS and IS, and the specificities of either.

One of the differences between (written) translation and (oral) interpreting is that the latter happens in real time, whereas the former stages interlinguistic encounters across longer temporal gaps. This has a strong impact on the nature of the language contact that happens in either type of activity. Indeed, interpreting is done in most cases in a near-instantaneous (consecutive interpreting) or instantaneous (simultaneous interpreting) manner, forcing the interpreter to produce the translation while concurrently having to manage the continuous flow of source-text input. This places tremendous cognitive demands on the interpreter, which may entail loss of accuracy and the occurrence of interference. Not surprisingly, cognitive approaches to language and memory have proved particularly helpful in the study of these spectacularly compressed forms of bilingual performance, while conversely, cognitive research can expect to learn much from the experimental observation of interpreting (see also Section 3.5. below).

The globally competitive and precarious nature of translation markets, and all too often, the poor management of multilingual communication projects, mean that much written translation now also has to happen under severe time pressure. However, compared to interpreting, there is generally far more time available for self-monitoring, self-editing and revision, and consulting various resources in the process. Crucially, the material fixation of written text allows for the possibility of contacts to be established with speakers/authors even from a remote past. It is the medium of writing that enabled, for example, Classical Latin to exert a delayed, but no less profound, influence on various European languages from the Renaissance onwards, which happened, in part, through translations from Latin into the vernaculars. In a slightly different way (but also along the diachronic axis), translation can put a language into contact with a former version of itself, in what is usually called (diachronic) intralingual translation. Examples include modernizations of the King James Bible, or translations of *Beowulf*, Chaucer, or Shakespeare into modern English. Inasmuch as the medium of writing (or print) enables translational contacts to happen across wide cultural and/or chronological divides, translators will often face problems of interlinguistic and intercultural interpretation. This has created a niche for hermeneutically inspired approaches within TS.

From the 1980s, TS scholars also started to take an interest in translation in a range of other media beyond the duality of orality and literacy. This began with work on dubbing and subtitling as forms of translation for TV and film. This kind of research has expanded in a major way, leading to the creation of a semi-autonomous subfield commonly labelled AVT (audiovisual translation). More recently, the digital wave has produced new previously unimaginable, technology-driven modes and forms of translation, both within and beyond AVT, indeed eroding the very distinction between AVT and 'general' TS, as their strong multimodal and transmedial dimension is challenging time-honored distinctions between word and image, between speech and writing, and perhaps in the longer run, even between man and machine. Such newcomers in 'translation land' include software localization, game localization, crowd translation, fan-subbing and other types of Internet translation, onscreen sign-language interpreting, audio description (for blind viewers), live subtitling (for deaf viewers) on the basis of automatic speech recognition, subtitling for opera and theatre, translation for museum audio guides, and so on. Even in mainstream translation, digital technology is a game-changer, offering working tools such as terminological databases, translation memories, and vastly improved automatic translation systems. The technology underlying these unprecedented man-machine interactions is entering everyday routines via our laptops and smartphones (see e.g. Google's recent Pixel Buds, designed to enable users to have 'direct' real-time conversations with people who speak a different language).

Needless to say, all these developments require a redefinition of what it means for texts, text users, and languages to come into 'contact' with each other.

3.5. Translator studies, and further perspectives

As was suggested by Chesterman (2009), TS has been moving away from 'translation studies' towards something like a more humanized 'translator studies', shifting research attention from texts and systems to the people (viewed both as individuals with particular personalities, life histories, ethical responsibilities, etc., and as members of social and

institutional structures) who take part in the complex process of producing, distributing, and using translations. This research trend implies an acknowledgment of the point made in Section 3.3 above, where the traditional source/text binary model, with the translator invisibly operating in the neutral middle between two systems, was criticized for being a reductive abstraction.

Examples include the ways in which critical discourse analysis and pragmatics have helped IS scholars to articulate how bilingual contacts in interpreting scenes always depend on dynamic interactions between all participants and interested parties involved in the process, and to draw attention to the difficulty of balancing the goals and concerns of individuals requiring linguistic mediation with those of the more powerful institution within which the bilingual proceedings take place (consider interpreting in medical or courtroom settings). Sociology, too, has helped to spotlight the translators with their personal life stories and in their relationships with other agents and social forces (e.g. Wolf and Fukari 2007). Pierre Bourdieu, and more recently Bruno Latour, have been major influences in TS. Latour's Actor-Network Theory invites us to attribute an agential role to technology in translation processes; given the developments sketched in Section 3.4 above, this proposition is worth exploring.

The increasing prominence of the cognitive paradigm in TS is another manifestation of growing attention being given to the human translator. The 'meeting' of source and target languages in the cognitive apparatus of individual translators and interpreters has been studied in process-oriented TS research for quite some time by means of TAPs (talk aloud protocols), and more recently, using more high-tech methodologies, such as eye-tracking and keystroke logging, separately or combined through triangulation. Corpus-based research involving advanced statistical techniques in TS is equally opening new avenues of investigation into the translator's mental 'black box' by complementing qualitative analyses with hard quantitative empirical evidence. This can provide new perspectives on the study of equivalence (perceived interlingual matchings or identifications), interference (source-language influence on the target text), and hypotheses of translation universals (interlingual effects proper to translation, yet irrespective of specific language pairs).

As Halverson (2014) has argued, a focus on individual translators and their cognition actually offers the attractive perspective of integrating the social world and the cognizing individual. The cognition of bilinguals is different from that of monolinguals by virtue of the copresence of different systems, whose activation must be controlled in order to avoid effects that would be socially undesirable in specific contexts, such as interference, code-switching, literal translations, etc. While we cannot in the current state of research determine how exactly (or greatly) translation differs from other bilingual activities, it stands to reason that much is to be gained by studying the two together within a cognitive framework. Along similar lines, Kruger (2017) has argued that studying translation together with other contact varieties from a combined cognitive and sociocognitive angle should permit a better understanding of the factors which shape translation as a special form of language contact, and which affect translation's potential for playing a role in language change.

(Socio)cognitively inspired work creates a framework for further convergence between CL and TS. Which of these two flags such research efforts prefer to sail under is of little (and indeed decreasing) importance, provided territorial claims and exclusionary effects are avoided. The same can be said about the converging work alluded to above

on translation policy (TS) and language policy (CL). However, this chapter has also deliberately dwelled on developments and conceptualizations in TS where connections with CL appear less plausible at first sight. This was done on the assumption that TS' expertise in several areas (including, among others, mediated and multilayered language contact, the impact of media, automatic translation and other communication technology, AVT, and the pragmatics and sociology of interpreting and translation) can enhance work in CL in ways which remain as yet unexplored. From the opposite perspective, one might surely envisage areas of cooperation between TS and CL wherein more of the existing research experience comes from the latter, such as the study of nonprofessional language brokering (e.g. in families, associations, companies, and other contexts) and language policy issues. CL's work on (multilingual) linguistic landscapes also reaches out to TS. If these opportunities materialize, the borderlines between CL and TS are bound to be even further eroded.

4. References

Angelelli, Claudia V.
2016 Bilingualism and multilingualism. In Claudia V. Angelelli & Brian James Baer (eds.), *Researching translation and interpreting*, 32–42. London & New York: Routledge.

Angelelli, Claudia V. & Brian James Baer (eds.)
2016 *Researching translation and interpreting*. London & New York: Routledge.

Baker, Mona
2006 *Translation and conflict: A narrative account*. London & New York: Routledge.

Benjamin, Walter
2012 [1923] On the task of the translator [transl. Harry Zohn]. In Lawrence Venuti (ed.), *The translation studies reader*, 3rd edn., 75–83. London & New York: Routledge.

Bermann, Sandra & Catherine Porter (eds.)
2014 *A companion to translation studies.* Chichester: Wiley-Blackwell.

Branchadell, Albert & Lovell Margaret West (eds.)
2005 *Less translated languages*. Amsterdam & Philadelphia: John Benjamins.

Chesterman, Andrew
2009 The name and nature of translator studies. *Hermes* 42. 13–22.

Cordingley, Anthony & Céline Frigau Manning (eds.)
2017 *Collaborative translation from the Renaissance to the digital age*. London: Bloomsbury.

Davis, Kathleen
2001 *Deconstruction and translation*. Manchester: St. Jerome.

D'hulst, Lieven & Yves Gambier (eds.)
2018 *A history of modern translation knowledge: Sources, concepts, effects*. Amsterdam & Philadelphia: John Benjamins.

Even-Zohar, Itamar
1990 *Polysystem studies* (special issue of *Poetics Today* 11[1]). Durham, NC: Duke University Press.

Gambier, Yves & Luc van Doorslaer (eds.)
2010–2013 *Handbook of translation studies*, 4 vols. Amsterdam & Philadelphia: John Benjamins.

Gambier, Yves & Luc van Doorslaer (eds.)
2016 *Border crossings: Translation studies and other disciplines.* Amsterdam & Philadelphia: John Benjamins.

González Núñez, Gabriel
2016 *Translating in linguistically diverse societies: Translation policy in the United Kingdom*. Amsterdam & Philadelphia: John Benjamins.
Halverson, Sandra L.
2014 Reorienting translation studies: Cognitive approaches and the centrality of the translator. In Juliane House (ed.), *Translation: A multidisciplinary approach*, 116–139. Houndmills, Basingstoke: Palgrave Macmillan.
Holmes, James S.
1988 [1972] *Translated! Papers on literary translation and translation studies*. Amsterdam: Rodopi.
Koller, Werner
1983 *Einführung in die Übersetzungswissenschaft*, 2nd edn. Heidelberg: Quelle & Meyer.
Kranich, Svenja
2014 Translations as a locus of language contact. In Juliane House (ed.), *Translation: A multidisciplinary approach*, 96–115. Houndmills, Basingstoke: Palgrave Macmillan.
Kranich, Svenja, Viktor Becher, Steffen Höder & Juliane House (eds.)
2011 *Multilingual discourse production: Diachronic and synchronic perspectives*. Amsterdam & Philadelphia: John Benjamins.
Kruger, Haidee
2017 Translation and/as language contact. Keynote address delivered at Translation in Transition 3 Conference, Ghent University, 14 July.
Malamatidou, Sofia
2016 Understanding translation as a site of language contact: The potential of the code-copying framework as a descriptive mechanism in translation studies. *Target* 28(3). 399–423.
Malmkjær, Kirsten & Kevin Windle (eds.)
2011 *The Oxford handbook of translation studies*. Oxford: Oxford University Press.
Meylaerts, Reine
2010 Multilingualism and translation. In Yves Gambier & Luc van Doorslaer (eds.), *Handbook of translation studies*, vol. 1, 227–230. Amsterdam & Philadelphia: John Benjamins.
Meylaerts, Reine & Theo du Plessis
2016 Multilingualism studies and translation studies: Still a long road ahead. In Yves Gambier & Luc van Doorslaer (eds.), *Border crossings: Translation studies and other disciplines*, 263–285. Amsterdam & Philadelphia: John Benjamins.
Munday, Jeremy
2012 *Introducing translation theories: Theories and applications*, 3rd edn. London & New York: Routledge.
Paulasto, Heli, Lea Meriläinen, Helka Riionheimo & Maria Kok (eds.)
2014 *Language contacts at the crossroads of disciplines*. Newcastle upon Tyne: Cambridge Scholars Publishing.
Pöchhacker, Franz
2016 *Introducing interpreting studies*, 2nd edn. Abingdon & New York: Routledge.
Pratt, Mary Louise
1991 Arts of the contact zone. *Profession* 91. 33–40.
Pym, Anthony
2008 Translation vs. language learning in international institutions: Explaining the diversity paradox. *Cultus* 1(1). 70–83.
Pym, Anthony
2014 *Exploring translation theories*, 2nd edn. London & New York: Routledge.
Schmidt, Heide (ed.)
1989 *Interferenz in der Translation*. Leipzig: Verlag Enzyklopädie.

Simon, Sherry
2013 Translation zone. In Yves Gambier & Luc van Doorslaer (eds.), *Handbook of translation studies*, vol. 4, 181–185. Amsterdam & Philadelphia: John Benjamins.
Venuti, Lawrence (ed.)
2012 *The translation studies reader*, 3rd edn. London & New York: Routledge.
Winford, Donald
2003 *An introduction to contact linguistics*. Malden, MA: Blackwell.
Wolf, Michaela & Alexandra Fukari (eds.)
2007 *Constructing a sociology of translation.* Amsterdam & Philadelphia: John Benjamins.

Dirk Delabastita, Namur & Leuven (Belgium)

Index

4-M(orpheme-type) model 19, 67, 153, 166

A

Aaron, Jessie Elana 482
Aasen, Ivar 387
Aboh, Enoch 26–7, 30
Abstand and *Ausbau* 4, 275, 347, 391–2, 394, 457, 542
abstract concepts 57–8, 60, 782
Abtahian, Maya R. 447
academia, languages of 399–400, 402, 541
Académie Française 359
academies, language 359, 360, 364, 736
Acadian French 481
accent
– 'foreign' 88, 201, 289, 335
– language attitudes 237, 238, 778–9
– pitch accent 476
– social psychology 778, 779
– standard 390
accommodation 118, 235, 249, 384, 385, 780
acoustic cues 92, 93
acoustic software 238
acquisition *see* first language acquisition; second language acquisition
acquisition planning 359–60, 363, 364, 721
acronyms 53, 338
activation, language 160, 161, 166
Actor-Network Theory 798
Adamou, Evangelia 20, 68, 473, 483, 645, 646
Adams, Michael 125, 126
Adamuti-Trache, Maria 192, 193
adaptation 203
additive bilingualism 407, 783–4
Adger, David 31
Adjunct Language Instruction 409
adlexification 54
adolescents 175–7, 178, 252 *see also* youth language
adoption 150 *see also* borrowing
adposition borrowing 18
adstrate influences 102, 108, 334
adult speakers 177–8, 180, 225
advanced learners 185
advocacy research 672
AfBo database 65
affective factors 235, 237, 529, 583
affixes 17, 19, 55, 56, 59, 646
Africa 320, 321, 322–6, 350, 442, 459–60, 549, 760 *see also specific countries*
African American English 285, 286
African American Vernacular English 337
African Americans 283, 285, 289, 312
Afrikaans 277, 325, 459, 462, 481
Afro-Hispanic languages of the Americas (AHLAs) 24, 27–34
age of speakers 172–84, 689
agency/agentivity
– Actor-Network Theory 798
– bilinguals as agents of lexical innovation 58
– contact-induced change 19, 25
– double agentivity 19
– emergence as trend 7
– epistemological shift towards 173, 314
– gender 246, 248
– language economics 708
– phonology 89–90, 92–3, 95–6
– political science 766
– Recipient Language (RL) versus Source Language (SL) 25–6
– turn towards 7, 583 *see also* individual
Agheyisi, Rebecca 237
Aghnihotri, Rama Kant 435
Ahlers, Jocelyn C. 683
Aiestaran, Jokin 436
Aikhenvald, Alexandra Y. 41, 44, 45, 47, 64, 448
Aikio, Marjut 248
Airey, John 541
Ajami 78, 79, 81, 84
Ajiboye, Esther 139
Akie people 461–2
Akkaya, Aslihan 690
Alaska 573, 594
Albania 376
Albanian 58, 65, 401, 552, 639
Albaugh, Ericka 350
Albirini, Abdulkafi 138
Aleut 69
Alexander, Charles H., Jr 509
Alexieva, Nevena 42
Algeo, John 44
Alhazmi, Albtool 426
Allen, Cynthia 25

https://doi.org/10.1515/9783110435351-065

Alsatian 116, 363
Altai 555
alternational code-switching 116, 161, 163
ALUPEC (Alfabeto Unificado para a Escrita do Cabo-verdiano) 83
Alzheimer's disease 225
Amdo Tibetan 564
American Community Survey 509
American Finnish 25
American Norwegian 64
American Sign Language 215, 217
American Swedish 64
Amerindian languages 699
Amharic 262, 435
Amit, Vered 311
Ammon, Ulrich 308, 361, 550
Anand, Pranav 639
anaphora 156
ancestral communication 461–2
ancestral languages 455–64 *see also* diaspora; heritage languages; indigenous languages
Anderson, Benedict 274, 298, 311, 445, 551, 731, 746, 767
Andrew, Patricia 172, 173, 178
Androutsopoulos, Jannis 84, 138, 141, 419, 420, 427, 491, 690, 755, 756
Angelelli, Claudia V. 791
Anglo-Americanism 236, 239
Anglocentricity 748
Angloromani 114, 148, 152–3
animacy 41, 55, 66
annotation tools 641–3
anonymity 301–3, 385
ANOVA (analysis of variance) 481–2, 567
Ansaldo, Umberto 198, 304, 588
anthropology 682–94
– collaborative 620
– colonialism 614, 618
– ethnography 587, 589, 689–90
– nexus analysis 583
– participant observation 613–25
– qualitative research 489
– social network analysis (SNA) 562
– speech communities 303, 309, 311, 313
– typology 556–7
Antonsen, Elmer 79
ANVIL 643
Aosta Valley 528
Appadurai, Arjun 142
apparent time studies 173, 179
Appel, Rene 45, 60
Applied Linguistics 174, 337, 435, 589, 747, 792
Arabic
– Arabic Pidgins 104–5
– borrowing 65, 151, 152, 154
– code-switching 426
– corpus linguistics 644
– diglossia 537
– domain analysis 539
– in France 352
– in Jerusalem 433
– nation states 344
– orthography 78, 81
– pragmatics 138
– qualitative data 495
– script choice 80
– in Tunisia 248
– in Turkey 349
Aragrande, Gaia 425
Aramaic 149, 260–1, 344
Arawakan 41, 65, 644
Ardal, Sten 663
Ardeşen Laz 66
areal linguistics 14, 40, 45, 46–7 *see also* geolinguistics
Arel, Dominique 526, 529
Arends, Jacques 102, 335
Argentina 176, 260
argot 325
Arnaut, Karel 165, 381, 592
Aromanian/Aromunian 65, 401
Aronin, Larissa 227, 230, 257, 267
Arop-Lokep 47
artefacts 582, 588
artificial languages 124–35, 543
artificial lexical networks 203
artificial miniature linguistic systems 226
Arvanitika 401
Asian Americans 289
Asian English 686
aspect 70, 130–1, 154, 200
assimilation 83–4, 249, 286, 288, 323, 344, 362, 374, 457, 462, 769, 773
Association for Language Awareness 223
Assyrian Empire 260
Asturias 528
asylum interviews 629–30
Atkinson, David 239, 423
Atlas of Pidgin and Creole Language Structures (APiCS) 70, 108
atlases, linguistic 524, 601
attentional resources 217

attitudes 234–45 *see also* ideologies
– anthropology 688
– to bi- and multi-lingualism 226, 352, 413, 423–4, 564
– to code-switching 164
– demolinguistic data 529
– domain analysis 546
– emergence as trend 7
– and language attrition 202, 203
– and language shift 445, 446–7
– migration 238, 239, 352
– to minority languages 781
– qualitative research 488
– surveys 503, 505, 508, 530
attractivity of forms 58
attributive markers 151
attrition 26, 29, 120, 162, 174–6, 179, 198–209, 423, 447, 757, 761 *see also* subtractive bilingualism
Audience Design 606
audio recordings 491, 494, 592, 601, 604, 607, 617, 642
audiobooks 749–50
audiovisual translation (AVT) 797
Auer, Peter 67, 141, 159, 164, 165, 172, 391, 392, 489, 540, 543
Ausgleich 364
Austin, Peter K. 307, 309, 310, 550
Australian Aboriginal languages 43, 260, 373, 644
Australian English 117, 138
Australian Kriol 43, 70
Austria 229, 445, 566, 633, 676, 683
authenticity 81, 239, 248, 301–3, 385, 604, 685, 769
autochthonous minority languages 393–4, 771 *see also* indigenous languages
autoethnography 619–21
auxiliaries 70, 114
available cases analysis 520
Avery, Peter 91
awareness, language 222–34, 412 *see also* metalinguistic awareness
Aymara 549
Azeri 80, 81

B

Baayen, R. Harald 483, 484
babbling 190
Bachelard, Gaston 280
back-flagging 161
Backhaus, Peter 432, 434, 437
Backus, Ad 198, 287
backwards diffusion 150, 155
Bahasa Melayu 351
Bahrick, Harry 193
Bailey, Benjamin 686, 690
Bajan 104
Baker, Colin 407, 719, 783
Baker, Lee D. 589
Baker, Mona 794
Bakhtin, Mikhail 240, 339, 577, 721, 726, 727, 745–6, 750
Bakker, Peter 19, 26, 34, 68, 69, 70, 71, 101, 104, 106, 162, 278, 684
Baldauf, Richard B. 364, 366, 699
Balkan Romani 18
Balkans 46, 150, 347, 645
Ballantyne, Davidson, and McIntyre v. Canada (1989) 378
Baltic German 538
Baltic states 46, 362–3, 540
Bamgbose, Ayo 79, 350
Bangkok 434
Baniwa 41
Barac, Raluca 225
Baraldi, Claudio 141
bare nouns 27, 28, 29–30
Barnes, Hillary 250
Bashkir 555
Basic English 128
Basic Varieties 193
Basque 66, 263, 347, 374, 410, 422, 424, 436, 528, 552, 639, 760, 768, 771
Batibo, Herman 79
Bauman, Richard 683
Bausani, Alessandro 126
BBC Voices 426
Becker, Alton L. 140, 603
Becker, Kara 603, 604, 607
behavioral skills 194
behavioral studies 211, 213, 503, 506
behaviorism 165
Bekker, Ian 236, 238
Belazi, Hedi 160, 161
Belfast 264, 564–5
Belgian Dutch 483
Belgium 348–9, 363, 392, 398, 437, 443–4, 501, 591, 630, 770
beliefs, language 358, 362, 366, 446, 529, 546, 688 *see also* attitudes
belief-sampling model 503

Belize 310, 447
Belize Creole 106, 447
Bell, Allan 606
Bell, Michael M. 321
Belling, Luc 265, 267, 424, 426
Belmont Report 669
Bemba 684
Bengali 351
Benor, Sarah Bunin 287
Ben-Rafael, Miriam 200, 203, 431
Bentz, Christian 19
Bequia 250, 251
Berber 352
Berbice Dutch Creole 70
Bergen 118, 310
Berger, Tilman 115
Bergmann, Christopher 200, 201
Bergs, Alexander 564
Berruto, Gaetano 550
Besnier, Niko 248, 673–4, 675
Betts, Dwayne 616
Beyogle, Richard 248
Bhabha, Homi K. 687
Bialystok, Ellen 199, 225, 442
bias (in research) 494, 508, 516, 530
Bible 390, 700, 796, 797
Bickel, Balthasar 14
Bickerton, Derek 26, 69, 101, 102, 106
bicultural accommodation 249
biculturalism 508
Bidese, Ermenegildo 78
bidialectalism 237
bidirectional influences 93, 95–6, 97, 199, 201, 215
big data 8, 509, 519, 530
bilingual education 407, 460, 593, 594, 686, 719, 722, 724–5, 744
bilingual first language acquisition (BFLA) 165, 166, 174, 186, 187, 188–9, 190, 191
Bilingual Interactive Activation Plus model (BIA+) 211, 213
Bilingual Language Interaction Network for Comprehension of Speech (BLINCS) 211, 213 bilingual language production model 166
bilingual lexical processing 212–13
bilingual optimization strategies 20
bilingualism *see also* code-switching; multilingualism; translanguaging; trilingualism
– acquisition 568–9, 640–1
– additive bilingualism 407, 783–4
– in adults 178
– agents of lexical innovation 58
– in artificial plus natural languages 124
– attrition 199 *see also* subtractive bilingualism
– bilingual communities and dialect variation 112
– bilingual first language acquisition (BFLA) 165, 166, 174, 186, 187, 188–9, 190, 191
– *bilinguismo assumido* ('conscious bilingualism') 83
– bimodal 215, 640
– and borrowing 66, 149
– children 174
– code-switching 165–6
– colingualism 323–4
– compound/subordinate bilingualism 44
– and converted languages 69
– corpus linguistics 640
– versus diglossia 537–8
– dynamic bilingualism 410, 413, 725
– economic value of languages 713
– educating with bilingualism 723–4
– education 406–18, 719, 723–5 *see also* bilingual education
– education for bilingualism 723–5
– elaborative bilingualism 45
– elite bilingualism 724
– family language policies 361
– folk bilingualism 724
– incipient bilingualism 257
– individual variation in bilingual lexical processing 210–22
– interpreting studies 796
– language attitudes 239
– linguistic dominance 89–90, 94
– versus multilingualism 257
– neuroimaging 662–5
– phonology 88
– political science 770
– and prestige 153
– pronunciation 201
– psycholinguistics 653–67
– social network analysis (SNA) 568–9
– speech-sign 215, 640
– subtractive bilingualism 407, 725, 783–4
– surveys 505, 508
– translation 793
– uneven development 190–1, 215, 259

– unstable bilingualism 26, 310, 461, 538, 770
biliteracy 77, 407
Bilkiire (Pidgin Fula) 70
bimodality 215, 216, 217, 640
bioprogram hypothesis 102, 103
biplots 476–8
Birdsong, David 33, 176, 185, 225, 656
biscriptality 79–80
Bislama 104, 250, 644
Black Twitter 312
Blackledge, Adrian 165, 339, 411, 582
Blainey, Darcie 609
Blair, Hugh E. 127
Blanc, Michel 226
Blank, Paula 58
Blanke, Detlev 126
Blasi, Damián E. 19, 27
blended languages 128, 338 *see also* mixed languages
blending of words 53
Block, David 582, 583
blogs 545
Blom, Jan-Petter 159, 163, 165, 401, 402, 540, 683
Blommaert, Jan 142, 143, 164, 165, 267, 287, 297, 299, 300, 333, 334, 337–8, 339, 362, 381, 419, 437, 445, 490, 577, 588, 590, 591, 592, 594, 626, 627, 629–30, 670, 685, 687, 688, 690
Bloomfield, Leonard 257, 258, 308
Blumenfeld, Henrike 216, 217
Blust, Robert 39
Boal, Augusto 620
boards/commissioners 736 *see also* academies, language
Boas, Franz 588–9, 615
body-part terms 39, 57
Bokhorst-Heng, Wendy 301
Boko 79
Bolivia 27, 590
Bollywood 646
Bonnici, Lisa 115
Booij, Geert E. 17, 65
Bora 65, 644
Borooah, Vani 713
borrowing 148–58 *see also* loanwords
– adposition borrowing 18
– alternative terms for 13
– artificial languages 130
– asymmetrical 15–18, 113, 473
– versus code-switching 60, 116, 148–9, 161–2
– cognitive processes 25, 153
– and colonialism 698
– constraints on 15–18, 19, 64, 68, 120–1, 149, 154–6
– and contact-induced change 13
– context of 65 *see also* cultural context
– continuum of borrowing-codeswitching 149
– cultural constraints on 19
– dating 53
– definition 148–9
– dialect variations 119
– differing cognitive processes 25
– direct versus mediated 25
– elicitation in research 605
– and gender 248
– heavy borrowing 152
– hierarchies 15–16, 18, 24–5, 42, 57, 63, 64, 90, 152–3, 155
– and language attrition 200
– MAT(ter) borrowing 13, 15–16, 40, 64–6, 68–9
– and the media 419, 420, 423
– mixed languages 68
– morphology 64–6
– motivations for 152–3
– for need or prestige 54
– nonce borrowing 162, 483
– PAT(tern) borrowing 13, 16, 40, 66, 69, 149
– phonology 89, 90
– semantics 40–3
– similarity to insertional code-switching 116
– social factors 66
– syntactic change 24–5
– varieties in contact 113–16
– word formation and vocabulary expansion 52–63
Bortoni-Ricardo, Stella M. 320, 568
Bosire, Mokaya 325
Bosnia 778
Bosnian 392
bottom-up language policies 361, 364, 390
Bou-Franch, Patricia 137
bound morphology 15–16, 18, 65, 70, 101, 106, 155, 160
boundaries 270–83
– anthropology 685, 687
– broadcasting 761

– education 413
– geolinguistic maps 531–2
– globalization 323
– ideologies 297, 298–300, 392
– linguistic landscapes 435
– media 419, 420
– speech communities 309
– standard language ideologies 392
– translation 795
bounded languages 300
Bourdieu, Pierre 138, 239, 299, 310, 313, 402, 633, 798
Bourhis, Richard Y. 431, 432, 436, 444, 456, 553, 555, 779, 781
Bowern, Claire 668, 676
Bowley, Arthur L. 515
boxplots 471–3
brain studies 4, 213, 662–5 *see also* psycholinguistics
branding 78
Branigan, Holly P. 166
Brazil 65, 94, 320, 373, 568
Brazilian Portuguese 201
Brenders, Pascal 214, 215
Brenzinger, Matthias 446, 454, 455, 462
Breton 263, 363, 457, 768
Brexit 264, 772
bridge system morphemes 67–8
bridging contexts 39
Briggs, Charles 607
Britain, David 117, 321
British Council 361
British English 114, 137, 254, 447, 484, 779
British National Corpus 640
broken transmission 148, 156
brokers, language 141, 799
Brookes, Heather 327
Brooks, Patricia 186
Broome Lugger Pidgin 106
Broth, Mathias 141
Brown, Penelope 102
Brunei 540
Brussels 115–16, 164, 192, 349, 433, 772
Brussels Dutch 115
Brussels French 116, 433
Bruthiaux, Paul 336
Brysbaert, Marc 658
Bucholtz, Mary 285, 291, 310, 313, 685
Budach, Gabriele 320
Bulgarian 18, 347
Bullock, Barbara E. 475
Bunčić, Daniel 79–80
Burke, Deborah M. 178
Busch, Brigitta 173, 275, 339, 577, 583, 627
butterfly effect 227
Bybee, Joan 178, 180
Byers-Heinlein, Krista 186

C

Cajun English 248, 250, 286
California 595
calquing 41, 43, 46, 53, 149
Calvet, Louis Jean 433, 696, 769
Cameron, Deborah 299, 420, 617, 626, 672
Cameroon 81, 104, 276–7, 311, 698
Cameroonian Pidgin 277
Camfranglais 277
Campbell, Lyle 25, 27, 41, 448
Canada 69, 193, 223, 278, 301, 348, 378, 380, 406, 407, 408, 432, 472, 481, 527, 528, 592, 603, 609, 709, 736 *see also* Québec; Toronto
Canadian English 484
Canagarajah, Suresh 140, 142, 173, 319, 322, 324, 337, 411, 591, 593, 595
Cantonese 41, 263
Canù, Silvia 83
Cape Dutch 481
Cape Verdean Creole 70, 83
Capotorti, Francesco 378
Cappadocian Greek 18
Caribbean Americans 291
Caribbean Creoles 104, 291, 684
Caribbean English 250
Caribbean Spanish 28
Carrington, Peter J. 563
cartography, linguistic 531–2
cascade model of attrition 204
case (grammatical) 14, 17, 63, 65, 66, 68–70, 149, 200
case studies 65, 77, 97, 114–15, 187, 236, 487, 645–6
Castañeda, Quetzil E. 620, 621
caste systems 236
Catalan 137–8, 275, 302, 303, 347, 552, 759
Catalonia 302, 303, 363, 374, 377, 402, 492, 528, 538, 770
categorical data 469, 474–5, 476, 482
categorical features 30
categorization 60, 151, 189, 237, 782
Catenaccio, Paola 425
Cavanaugh, Jillian R. 446–7

Celtic 57
Cenoz, Jasone 257, 407, 408, 434, 436
censorship 426
census analysis 236, 351, 443–4, 469, 501, 515, 525–7, 529–30, 580–1, 617
Centre international de Recherche sur l'Aménagement linguistique 3
Chambers, J. K. 392
Chamorro 699
Chana, Urmi 164
Charles V, Emperor 261, 266
Charter for Regional or Minority Languages 773
Chaudenson, Robert 103
Chaudron, Craig 225
Chen, Peiyao 217
Chengappa, Shyamala 655
Cherciov, Mirela 203
Cheshire, Jenny 175, 177, 606, 609
Chesterfield, Kathleen 191
Chesterfield, Ray 191
Chesterman, Andrew 797
Chicano Civil Rights movement 275
Chicano English 286, 289
Chierchia, Gennaro 29
child language acquisition *see* first language acquisition
Child Language Data Exchange System (CHILDES) 187, 641, 643
children 173–5, 188–93, 200, 212, 225, 408, 569, 591, 640–1, 654, 689
Chiluwa, Innocent 139, 351
China 492, 750
Chinchano Spanish 27–34
Chinese 41, 81, 113, 139, 263, 301, 345, 351, 352, 446, 566, 644 *see also specific Chinese languages*
Chinese Americans 288
Chinese communities 566–8, 569
Chinook Jargon 105
Chipewyan 448
chi-square 474–5, 476
Choi, Jinny K. 474–5, 476
Chomsky, Noam 30, 161, 165, 258, 643, 778
Chorote 260
Chota Valley Spanish 27–34
Chow, Rey 744
Christoffels, Ingrid K. 661
chronotopes 577
Churchill, Stacy 374
Cichocki, Wladyslaw 474, 476
Cimbrian 78
cities, multilingual 263–4 *see also* superdiversity; urban areas
citizenship 301, 371, 378, 381, 736, 772–5
classifiers 14, 55
clausal syntax 107, 150
Clements, Joseph 65
Clifford, James 619
clipping 53
cluster analysis 468, 480–1
Clyne, Michael 1, 3, 137, 159, 160, 162, 166, 203, 286, 287, 326, 442, 443, 445, 448, 456, 553, 555, 587
co-activation 166, 202, 210, 212, 214, 215, 216, 217
Cobarrubias, Juan 362
codabillity, language 506
code alternation 323
code-copying 13
codemeshing 140, 322, 324
code-mixing 58, 540
code-switching 159–71
– alternational code-switching 116, 161, 163
– anthropology 683
– versus borrowing 60, 116, 148–9, 161–2
– children 174
– code choice 536, 539
– computer-mediated discourse 138
– constraints on 159, 160–1, 166
– continuum of borrowing-codeswitching 149
– corpus linguistics 643, 645, 646
– crossing, passing, refusal 322–3
– dialect variations 119
– domains 400, 539–40
– versus dynamic bilingualism 410
– education 410–11
– ethnography 159, 164, 166
– ideologies 539
– indexing identity 138, 164, 540
– inflectional morphology 67
– informal written corpora 639
– insertional code-switching 67–8, 116, 141, 161, 162, 163, 190, 338, 646
– interviews (research method) 606, 609
– and language attrition 202
– lexical borrowing 60
– linguistic boundaries 279
– linguistic landscapes 435
– literary studies 747, 749
– and the media 419, 423–4

– metaphorical code-switching 164, 401, 402, 489, 683
– methods for social media research 425–6
– methods for studying 544–5
– Minimalist Program 643
– mixed languages 68
– and morphological borrowing 19
– multilingual contexts 259
– older people 180
– power dynamics 630
– pragmatics 137
– psycholinguistic research 660
– qualitative research 489
– researching acquisition of 569
– situational code-switching 164, 489
– speech communities 311
– as strategy 164–5
– structural similarity 65
– translanguaging 140, 413
– Twitter 425
– urban-rural contact 322–3
– varieties in contact 116–18
– in work teams 425
codification 346, 347, 387–9, 542, 543–4, 594, 697–8, 700, 758 *see also* corpus planning
coding of data 494, 641
cognates 211, 212, 213, 214–15, 659
cognitive processes
– adults' 178
– bilingualism 211
– borrowing 25, 153
– cognitive mobility 774
– constraints on formation of creoles 103
– cross-language activation 215–18
– declining through age 179, 225
– demolinguistic data 529
– interpreting studies 796
– and language attrition 200, 201
– metacognition 230
– metalinguistic awareness 224, 225, 226
– overt subject pronouns 32
– pidgin genesis 102
– retrieval for survey purposes 505–6
– semantic change 44–6, 48
– translation 798
– word formation and vocabulary expansion 60
coinages 42, 56, 59
co-lexification 48
colingualism 322, 323–4, 326
collaborative corpus linguistics 641
collaborative research 619–21, 671 *see also* interdisciplinarity
collocational context 59–60, 639
colloquial language 56
colonialism 695–706
– anthropology 614, 615, 618
– colingualism 324
– colonial languages 45
– colonial studies 695–706
– contact-induced syntactic change 28
– corpus linguistics 644
– educational sciences 721
– empire, and multilingualism 261
– and English 236, 325, 593
– ideologies 697–8, 700, 703
– indigenous languages 373, 462
– language attitudes 239
– and language shift 442, 447
– linguistic boundaries 276
– and linguistic nationalism 350–1
– and migration 333
– multilingual contexts 260
– and named languages 727
– one nation-one language 300–1
– and orthography 79, 83
– participant observation 615–16
– pidgins and creoles 28, 71, 105, 543, 697, 701
– pluricentric languages 392
– premodern empires 344–5
– translanguaging 727
– varieties in contact 113
color terms 57
Combs, Mary Carol 590
Comeau, Liane 166
commodification 716
Common European Framework of Reference (Council of Europe 2001) 229
Common Underlying Conceptual Base 227
Common Underlying Proficiency 227
Communication Accommodation Theory 159, 163, 780–1
communication studies 754–65
communicative competence 309, 407
communicative language teaching 411
communicative moves 141
communicative needs 227–8
communities of practice 264–6, 312, 313–14, 564, 575, 690
community, concepts of 311–12
community language maintenance 248

Community-Based Language Research (CBLR) 673
comparative linguistics 152, 433
Comparative Method 39–40
comparative structural approaches 26
compartmentalization 154
compensatory strategies 200
competence 179, 267, 309, 407, 545, 722
competition/adaptation 200, 201, 203, 217
Complementizer Phrases 67
complexity, linguistic 176, 227, 266, 277
composite matrix languages 68–9
compounding 107
compound/subordinate bilingualism 44
comprehension, acquisition of 186
computational methods 425, 641
computational phylogenetics 72
Computer Assisted Personal Interviewing (CAPI) 507
Computer Assisted Telephone Interview (CATI) 507, 526
computer-mediated discourse 138, 419–20, 690 *see also* internet-based media; social media
Comrie, Bernard 125
conceptual boundaries 271, 278
conceptual expansion 42, 47, 60
conceptual ontologies 39
conceptual reorganization 41
conceptual universality 60
concordances 639
conditional inference trees 468, 483–4
conditioning factors 545–6
conflict 5, 239, 274, 279, 364, 436, 625–6, 713, 731, 737, 794
conflict resolution 216
congruence 65, 68
congruent lexicalization 160, 161, 163, 166
conjunctions 16, 20, 155
Conklin, Nancy R. 444
connectives 155
connotational meaning 60, 114
consciousness 224, 250, 361
consent 491, 607, 669, 675, 676
conservatism 56–7, 177, 179, 249, 447
constructed languages 124–35
construction, language 756 *see also* maintenance of languages
constructivism 304
contemporary urban vernaculars 178, 324
Content and Language Integrated Learning (CLIL) 408, 409–10, 724
Content Based Instruction (CBI) 406, 408, 409
contextual inflection 17
contextualization 164, 543
contextualization cues 137, 323, 489, 493, 496
contingency tables 474–5, 476
continuous data 469
continuum of borrowing-codeswitching 149
continuum of contact (casual − intense) 18, 24–5, 54, 64, 97, 155, 202, 684, 697
continuum of dialects 391–2, 393 *see also* *Abstand* and *Ausbau*
continuum of language dominance 549 c
ontinuum of multilingualism 259, 309, 310
continuum of pidgins-creoles 100, 334
continuum of rural-urban areas 320–1, 322, 324
Contrastive Analysis 229
Contrastive Hierarchy 6
contrastive linguistics 137
convergence
– anthropology 684
– borrowing 149, 150
– code-switching 162, 163
– corpus linguistics 645
– phonological change 93, 96
– semantic change 45, 46–7
– social psychology 780–1
– standardization 384, 386–9, 394
Conversation Analysis (CA) 141, 159, 165, 489, 494
converted languages 69
Cook, Vivian 140, 198, 226, 229, 411, 448
Cook-Gumperz, Jenny 311
Cooper, Robert L. 237, 358, 433
coordinate bilingualism 44
Cop, Uschi 213
Copp Jinkerson, Alicia 411
Copp Mökkönen, Alicia 412
copulas 102–3, 107
copying 13, 43 *see also* borrowing
Cormack, Mike 756, 761
Corngold, Stanley 743, 745
Cornish 346, 552
Coronel-Molina, Serafín M. 592
corpus data
– borrowing 20
– corpus linguistics 425, 632–4, 638–53, 798
– corpus methods of research 469
– dialect variations 119

– emergence as trend 8
– pragmatics 137–8
corpus planning 346, 347, 349–51, 359, 362, 363–4, 542, 721
correlation coefficients 474
correspondence analysis 475–8, 483, 484
Corsetti, Renato 131
Corsica 299
Corsican 299, 363
Cortès-Colomé, Montserrat 239
cosmopolitanism 325, 327
Costa, Albert 659, 660, 662
cot-caught merger 115
Cots, Josep 223
Cotter, Colleen 755
Coulmas, Florian 77, 79, 80–2, 84–5, 314, 432, 455, 456
Council of Europe 229, 363, 379, 393, 409, 631, 773
count/mass nouns 30
Coupland, Nikolas 178, 179, 297, 319, 384, 426, 434, 490, 755
covariance 469
covert norms 310, 314
covert policies 358, 361, 362
covert prestige 114, 327
covert values 364
Coyle, Do 409
Cramer's V 475
Crawhall, Nigel 462
creativity 53, 60–1, 139, 176, 200, 226, 420
Cree 19, 69, 278, 549
Creese, Angela 165, 339, 411, 582
creoles
– adstrate influences 102, 108, 334
– anthropology 684
– artificial languages as 131
– colonialism 28, 71, 105, 543, 697, 701
– constraints on formation of 103
– contact-induced change 19, 26
– continuum of pidgins-creoles 100, 334
– Creole Debate 26
– Creole Prototype 70, 71
– creolization 100–12, 310, 344, 543, 684
– Decreolization Hypothesis 28, 30
– domains 538
– ecological factors 27
– exceptionalism 71, 701
– formation of 102–3
– gender stratification 250
– and globalization 335
– intermediate creoles 26, 104, 105
– linguistic boundaries 278
– linguistic properties 106–8
– and the media 424
– Monogenesis Hypothesis 28
– morphology 69–71
– origin of term 100
– orthographic reforms 83
– and speech communities 310
– substrate influence 13, 27, 45, 58, 101, 102, 103, 106, 334, 447
– superstrate influence 101, 102, 103, 334
– syntactic change 26
– typology 26, 27, 69, 71, 103, 107, 701
creoloids 334
critical discourse analysis (CDA) 240, 425, 492, 627, 798
critical period hypothesis 174, 176, 178, 656
critical social sciences 359
critical sociolinguistics 627, 631
Critical Theory 744
Croatia 344
Croatian 348, 392, 473, 778
Croft, William 202, 683
cross-generational comparisons 290
cross-generational contact-induced change 29
crossing 177, 279, 322–3, 493, 686
cross-linguistic influence 137, 165, 186–7, 189, 199, 210–22, 226, 227–9
cross-sectional studies 187, 193
crowdsourcing data 641
Crowley, Terry 101
Cru, Josep 299, 426
Cubeo 41
cultivation, language 358
cultural anthropology 562
cultural context
– anthropology 685
– and borrowing 53–4, 57–8, 59
– cognitive processes of multilingualism 45
– content and language integrated learning (CLIL) 410
– cultural capital 419, 689
– cultural contact 47, 114, 120, 247, 684
– cultural frame switching 508
– cultural loans 54, 114
– cultural mediation 574
– cultural pluralism 275
– education 412
– and ethnicity 285–7
– ethnographic research 589
– gender 247

– intercultural communication 4, 42, 136, 141, 249
– interviews (research method) 607
– language economics 709
– language ideology 297, 300–1
– language shift 445
– lexical borrowing 56, 59
– linguistic boundaries 276
– literary studies 747
– localization 421
– media 419, 755
– mixed languages 325
– multiculturalism 264, 285–6, 375, 437, 542, 749, 771
– pragmatics 137, 138
– race and ethnicity 284–5
– revitalization of languages 769
– semantic borrowing 42, 45
– and semantic change 39, 47
– social psychology 779
– surveys 506, 508
– translation 793, 794
– writing reforms 81
Cummins, Jim 227, 406, 411, 412, 414, 415, 726
Curtin, Melissa L. 594
Cushing, Frank Hamilton 614, 616
Cyrillic script 80, 81
Czaykowska-Higgins, Ewa 673
Czech 113, 114, 115, 117, 118, 119, 349

D

Dahrendorf, Ralf 774–5
Dakar 433
D'Angelo, Alessio 565
Danish 386, 387, 388, 390, 391, 392, 400, 401, 495
Dankova, Natalia 130
Darquennes, Jeroen 359, 363, 364, 374, 393, 412, 426, 436, 443, 556, 557, 625
Dasher, Richard B. 39
data coding 494, 641
data matrix presentations 470–1, 475
data mining 530 *see also* big data
David Maya, Khemlani 180
Davis, Bronwyn 629
d-BIA (developmental Bilingual Interactive Activation model) 211, 213
de Bot, Kees 174, 177, 178, 179, 180, 225, 425, 426
de Bres, Julia 265, 267, 424, 426
de Granda, Germán 28
de Groot, Annette M. B. 210, 211, 212, 659, 661
De Houwer, Annick 165, 174, 188, 189, 190, 191, 193
de Pietro, Jean-François 412
de Saint Georges, Ingrid 574, 583
De Sutter, Gert 483
de Varennes, Fernand 371, 376, 378, 381, 733, 735, 738
De Vries, John 524–5
death, language 180, 462, 683, 696 *see also* endangered languages
DeCamp, David 479
Declaration of Helsinki 669
decolonization 588, 618, 619–20, 727
decreolization 28, 30
definite articles 149, 151, 154
DeGraff, Michel 69, 70, 71, 100, 701
deixis 154, 156
Delaere, Isabelle 483
Delaware Pidgin 102, 105
delegitimization 347
Delhi 320, 435
Dell'Aquila, Vittoria 444, 524, 528, 530, 531, 532, 550
demasking techniques 658
dementia 179
Democratic Republic of the Congo (DRC) 324, 435
demographic studies 236, 503
demolinguistic data 523–35
denaturalization 334–6
Dench, Alan 71
dendrograms 480–1
Denmark 348, 402, 541
Deo, Ashwini 43
derivational affixes 17, 53, 59, 63, 64–5, 101
Derrida, Jacques 744, 795
descriptive statistics 471–81
de-standardization 400
Determiner Phrase (DP) agreement 28, 30
determiners 67, 161, 200
Detroit 313, 564
Deumert, Ana 320, 364, 385, 386, 481, 701
Deutsch, Karl W. 766
Deutscher, Guy 685
development economics 713
developmental bilingual education programs/ dual language 725

developmental Bilingual Interactive Activation model (d-BIA) 211, 213
developmental perspectives 177, 185, 186, 188, 189, 203, 226, 569
Di Carlo, Pierpaolo 311
Di Sciullo, Anne-Maria 159, 160
diachronic approaches
– artificial languages 130
– code-switching and borrowing 60, 162
– domains 540–1, 545
– language acquisition 186–7
– language policy and planning (LPP) 366
– and the media 424, 426
diachrony of meaning 39–40, 48
diacritics 78, 85, 758
dialect contact 112–23, 162, 278, 392
dialect continuum 391–2, 393 *see also* *Abstand* and *Ausbau*
dialect geography 601
dialect levelling 116, 117–18, 162, 322
dialect variations *see also* minority language groups; varieties in contact
– interdialect forms 116–18, 322
– and language ideology 301
– language politics 347, 364
– and the media 421–2, 756
– Norway 387, 390
– phonology 289
– social networks 540
– status of 300
– urban-rural contact 321, 327
dialectology 239, 524, 601
dialect/standard constellations 391–2
diaspora 130, 248, 261, 265, 266, 311, 423, 426, 455, 458–61, 550, 593, 606 *see also* heritage languages
Díaz Noci, Javier 422
Dick, Hillary Parsons 688
dictionaries 350, 364, 387, 758
Diebold, A. Richard 257
diffusion of change 149, 150, 155, 204, 252, 565
diffusion tensor imaging (DTI) 664–5
digital technology *see also* computer-mediated discourse; internet-based media; social media
– media linguistics 425–6
– multidisciplinarity 9–10
– multilingual communities of practice 265
– networked multilingualism 84
– online communities 312
– pragmatics of digitally-mediated interaction 141
diglossia 112, 153, 263, 279, 309, 326, 391, 488, 537–44, 683, 724, 769–70
digraphia 79–80
Dijkstra, Ton 166, 211, 212, 213, 214, 215, 658
DIMEAN (*diskurslinguistische Mehr-Ebenen-Analyse*) 632–3
Dinkin, Aaron 290
diphthongization 115
directness 137
disability studies 618
disambiguation tasks 655
discourse analysis 138, 139, 240, 492, 574, 576, 583, 594, 625–38, 754
discourse communities 312
discourse linguistic multilevel analysis 632
discourse markers 27, 161
discourse phenomena 45, 103, 136 *see also* pragmatics
discourse planning 360
discourse variables 606
discourse-historical approach (DHA) 627
discourse-managing elements 45
discourse-oriented approaches 239, 240, 359, 360
discourse-situational factors 166
discrepancy-arousal theory 780
discrimination 363, 378, 380, 419, 457, 635, 676, 709, 713, 733
discursive order 627, 628
disfluency 200
distancing 248, 271, 279, 629
distributed feature model 211
distributed language use (4E model) 746
divergence 45, 163, 386–9, 394, 434
Dixon, Robert M. W. 64
Djité, Paulin 79
Docherty, Gerard J. 89
documentation, language 613, 682, 697–8, 702
domain analysis 536–49
domain loss 399–400, 536, 541
domains 39, 152, 201, 279, 309, 361, 397–405, 448, 536–49, 710
domestic work contexts 104–5
Dominican Americans 686
Donohue, Mark 14
doppel avoidance 45

Dorian, Nancy C. 3, 249, 250, 445, 448, 454, 474, 683
Dörnyei, Zoltan 518, 519
Dothraki 124
double agentivity 19
doublets 57
Doughty, Catherine J. 45
Dragojevic, Marko 237, 238, 240, 778, 780, 781
dramatization 279
Drinka, Bridget 114
dual scaling 475
Duala 698
Duarte, Joana 165
dubbing 422, 797
Dubois, Sylvie 248, 288
Duchêne, Alexandre 526, 527, 529, 626, 627, 631
Dueber, Dagmar 325
Dupont, Véronique 320
Durkin, Philip 53
Dutch 84–5, 93, 94–5, 116, 137, 201, 212, 214, 392, 433, 483, 609, 630, 644
dynamic bilingualism 410, 413, 725
Dynamic Model of Multilingualism (DMM) 222, 227–30
Dynamic Systems or Complexity Theory (DSCT) 227–30
Dynamic Systems Theory 178, 203, 227–30
Dynamics of multilingualism with English (DyME) 228, 229

E

Early Second Language Acquisition (ESLA) 189–90, 191
earning differentials 709–10, 713
ear-voice-span 661
East African Gujerati 460
East African Swahili 549
East Sutherland Gaelic 445
Eastern Ijo 70
Eastern Maroon Creoles (EMC) 102–3, 107
EasyAlign 643
Eckert, Penelope 173, 175, 179, 235, 239, 240, 246, 287, 308, 313, 315, 564, 613, 669, 677
Eckhardt, Regine 40
Eckman, Fred 97
Eco, Umberto 128
ecolinguistics 10, 364–6, 594
ecologies, linguistic 27, 311, 442, 444, 551, 553, 593–4, 696, 701, 737, 746
economics of language 707–18
ecosystems, linguistic 365–6, 758–9
Ecuador 27, 278, 555, 593
Eddington, David 484
education 406–18
– bilingual education 407, 460, 593, 594, 686, 719, 722, 724–5, 744
– default writing systems 79
– ethnography 589, 593–4
– ideologies 362
– immersion programs 225, 406, 407, 408–9, 410, 724
– instructed versus uninstructed acquisition 7
– language policy and planning (LPP) 363
– language politics 360, 361, 774
– medium of instruction 266, 376, 381, 399, 400
– minority groups 376, 771
– minority language rights 379
– modernism 721–5
– nation states 343, 345, 346
– nexus analysis 582
– two-way/dual language programs 407
educational linguistics 577, 589
educational sciences 719–30
Edwards, John 338, 364, 444, 551, 553, 555, 556, 557
Egbert, Maria 141
ego networks 566–7
Ehala, Martin 540, 543, 545, 546, 550
elaborated/expanded pidgins 104
elaboration 389–90
elaborative bilingualism 45
ELAN 643
elderly people 178–80, 185, 225
ELDIA 550
electroencephalography (EEG) 663
electrophysiological studies 211, 213, 217
elite bilingualism 724
Ellison, T. Mark 45
Elšik, Viktor 16, 153, 155
Embedded Language (EL) 161, 162
embedded languages 148
embodiment 139, 577, 582, 583, 614, 686, 746
emergent properties 228–9
emic perspectives 164, 589, 594, 603, 609, 613, 616, 690
emoticons 338, 495

empire, and multilingualism 260–1, 344–5, 348, 695
empowering research 672–3
empowerment discourses 9
encyclopedic views of meaning 39, 47
endangered languages
– age of speakers 179–80
– anthropology 683
– corpus linguistics 640
– domains 399, 536
– ecolinguistic crisis 10
– ethnographic research 591
– ideology 688
– language maintenance 454, 457
– language politics 362, 769
– language shift 443, 444, 447
– linguistic landscapes 436
– and the media 755, 756
– nation states 350
– and research ethics 670, 673
– social network analysis (SNA) 568
– typology 550, 555
Enfield, Nicholas J. 39, 40, 41, 46, 48
Engelberg, Stefan 699, 701
English
– in Africa 83, 549
– bilingualism 201
– borrowing 150, 153, 248
– bridge system morphemes 67–8
– in Brussels 433
– calques from 53
– in Canada 432, 784
– and codification of other languages 388, 389
– colonialism 236, 325, 702
– computer-mediated discourse 138
– corpus linguistics 644
– creoles 105, 106
– domains 398, 540
– English-only ideology in education 411
– in Eritrea 83
– as foreign language 712
– fortis-lenis systems 91
– as global language 238, 266, 338, 352, 372, 422, 435, 593
– and globalization 266
– in India 381, 689
– infant acquisition of 186
– instrumental reasons for learning 236
– as international language 236, 335, 399, 541, 712, 738
– and the internet 759
– and Japanese 114
– in Jerusalem 432–3
– in Korea 423
– language economics 712, 713, 715
– as language of academia 399–400, 402, 541
– languages of wider communication (LWCs) 442
– as a lingua franca 238, 263, 266, 325, 335, 338, 737, 738, 766
– linguistic justice 544
– linguistic landscapes 435
– loanwords 55, 56, 58, 59, 64, 88, 90
– in Malaysia 351
– and the media 422, 423
– medium of instruction 400
– mixed languages 337
– Old English 44, 79
– in Pakistan 351
– in Papua New Guinea 260
– phonology 94–6
– pragmatics 137–8
– proficiency of multilinguals 214
– in Scotland 761
– script choice 79
– in Singapore 263–4
– social values 248
– in South Africa 277, 371, 459
– and structural (mediated) transfer 26
– stylistic additions 57
– in Sweden 594
– syntactic variation 484
– trading language 713
– translation 750
– in the US 275, 527, 688
– World Englishes 336, 447, 484
– writing systems 77
English as a Lingua Franca (ELF) 335–6
entailed meanings 39, 40, 43
entextualization 138, 631
environmental factors 27, 179–80, 237–8
epilinguistics 271
epistemic violence 616
Epps, Patience 19, 34, 43, 44, 46, 47
equality 769 *see also* inequality/injustice; justice, linguistic
equivalence, linguistic 792–4, 798
Equivalence Constraint 159, 160
Erickson, Frederick 589
Eriksen, Thomas 373, 374
Eritrea 83
Errington, Joseph 683, 688, 695, 696, 698

errors 504–7, 519–21, 530
Esau, Katrina 462, 463
Escobar, Anna María 2
Esperanto 124, 125, 126, 128–9, 130
essentialism 267, 303, 315, 687, 689
Estonian 362, 538, 541
ethics 8, 491, 607, 616, 618, 641, 668–81, 716, 795
Ethiopia 39, 46, 262, 435
Ethiopic 83
ethnicity
– anthropology 686
– ethnic (urban) minorities 375
– ethnic cleansings 378
– and gender 247, 250, 252, 286
– and identity 283–96
– interethnic communication 104–5
– language attitudes 235
– language shift 443
– multiethnic communities 260–3, 264, 265, 275, 276, 283–4, 325
– multiethnolects 175, 177, 178, 286–7
– multilingual contexts 261
– social boundaries 277
– social networks 566–7
– studying 284–6
– style-shifting 606
– urban-rural contact 325
– within-group differences 290–1
ethnocultural conflict 377
Ethnographic Linguistic Landscape Approach (ELLA) 437
ethnography 587–600
– and anthropology 587, 589, 689–90
– autoethnography 619–21
– code-switching 159, 164, 166
– collaborative ethnography 619–21
– and colonial studies 697
– of communication 309
– and discourse analysis 634–5
– epistemic and disciplinary foundations 588–90
– fieldwork methods 489–92
– and gender 248
– interactionist research approaches 488, 489
– interviews (research method) 491, 602
– language planning 359
– language shift 445
– linguistic landscapes 434
– move towards the individual as focus 314
– multi-site ethnography 619–21
– nexus analysis 575, 578, 579, 580, 581, 584
– participant observation 613–25
– performativity 620–1
– speech communities 308, 311
– way of being 588
ethnolects 164, 286–7
ethnolinguistic discrimination 709, 713
ethnolinguistic repertoires 287
ethnolinguistic variation 235, 237, 284, 314, 350, 351, 531
ethnolinguistic vitality 203, 444, 456, 546, 553
ethnomethodology 165
Eurobarometer 386, 502, 503
Eurocentricity 100, 267 *see also* Western bias
European Charter for Regional or Minority Languages 363, 371, 379, 393, 733, 734, 736, 737
European Commission 409
European Parliament Proceedings Parallel Corpus 641
European Union (EU) 263, 352, 361, 363, 444, 556, 590, 688, 734, 738, 761, 772
Evans, Nicholas 17, 39, 260, 461
Éveil aux langues (Evlang) 223, 412
Evenki 66
Event Related Potentials (ERP) studies 213, 663
Even-Zohar, Itamar 793
Everett, Daniel L. 148, 156
evidential categories 44, 46
evolution of human language 156
executive control 153, 155, 215–18, 225
EXMARaLDA 643
exogamy, linguistic 19, 34, 260, 277, 459, 460, 683
exoglossic standards 386, 387, 390, 391
Expanded Graded Intergenerational Disruption Scale (EGIDS) 444, 456
expanded pidgins 105
expectancy bias theory 782–3
experimental research 468–9, 481–2, 639
expert choice sampling 518
Extended Projection Principle (EPP) 30
Extra, Guus 375, 555
extralinguistic framing 6, 18, 118, 120–1, 483, 524, 555, 642, 698
eye movement studies 213, 216, 217, 656–7, 798

F

Facebook 138, 265, 337, 423, 425, 426, 495, 518
factor-analysis 237, 468
Fagyal, Zsuzsanna 564
FAIR data (Findable, Accessible, Interoperable and Reusable) 646
Fairclough, Norman 240
Faist, Thomas 333
false friends 214, 215
family language policies 361
Fanakalo (Pidgin Zulu) 70, 106
fangyan 345
fantasy languages 129
Fasold, Ralph 326, 479, 480
faux fonts 84–5
Fayzrakhmanova, Yu 423
features 30–2, 90, 91, 150, 211, 251
feminist theory 362, 744
Ferguson, Charles F. 139, 279, 309, 488, 537, 550, 552–3, 683
Ferreira, Victor S. 661
Fessi 327
Feyaerts, Kurt 139
fictional languages 124
Field, Frederic W. 15
Field, Margaret C. 447
Filipino communities 568
Findling, Joav 399
Finkbeiner, Matthew 211
Finland 262, 363, 380, 411, 527, 577, 756, 757, 760, 768
Finnish 25, 262, 275, 393, 394, 527, 644, 768
Finnmark 374
first language acquisition
– bilingual first language acquisition (BFLA) 165, 166, 174, 186, 187, 188–9, 190, 191
– children 174
– code-switching 165–6
– constructed languages 131
– contact-induced syntactic change 29
– corpus data 640–1
– psycholinguistics 654–5
– uninstructed language acquisition 186, 188–9
First Nations languages 528
Fischer, Andreas 53
Fishman, Joshua A. 10, 67, 77, 80, 139, 163, 202, 237, 259, 263, 275, 279, 309, 314, 346, 359, 363, 398, 399, 401, 402, 441, 442, 443, 444, 456, 488, 490, 537, 538, 539, 540, 541, 543, 555, 588, 721, 724–5, 756, 769
flanker tasks 217
Flege, James E. 201
Fløgstad, Guro 176
Flores, Nelson 406, 413, 414
fluency 200, 201, 259
Flynn, Suzanne 227
focusing 162, 279
folk bilingualism 724
folk linguistics 239, 309
footing 279
foreign language learning 192, 724 see also second language acquisition
Foreigner Talk hypothesis 335
foreignization 790, 794
foreignness, perceived 58–9
formalism 626
formality 137, 138, 153, 401
form-meaning relationship 38, 40–3
formulaic expressions 189, 192
fortis-lenis systems 89, 91
fossilization phenomena 31
Foucault, Michel 626–7, 628, 630, 632, 634, 721, 726–7
Foundation for International Development Study and Research (FERDI) 273
fractals 272, 297
frames 39, 164
Framework Convention for the Protection of National Minorities 363, 371, 379, 734
François, Alexandre 45, 47
Free Morpheme Constraint 159, 160, 161
Free your stuff Luxembourg 265, 267
Freisner, Michael 290
French
– Acadian French 481
– in Belgium 348–9
– bilingualism 201
– borrowing 26, 55, 56–7, 59, 248
– in Brussels 116, 433
– in Canada 348, 408, 432, 481, 779, 784
– as colonial language 702
– corpus linguistics 644
– creoles 106
– French-Cree 19
– and globalization 352
– international language 399
– languages of wider communication (LWCs) 442

– linguistic landscapes 433, 435
– in Luxembourg 265
– and the media 423
– mixed languages 278
– morphology 69
– Norman French 17
– phonology 89
– pragmatics 137, 138
– Prince Edward Island French 26
– script choice 82
– in Senegal 265
– social values 248, 249
– in Switzerland 712
– trilingualism 214
– in Tunisia 248
– uniformization 768
– in the US 552
– voicing distinctions 91, 93
French Guiana 311
frequency distributions 471, 474–5
Friesland 363, 528
Friulian 552
frozen actions 582
Fuller, Janet M. 686
function words, borrowability of 16
Functional Head Constraint 160
functional magnetic resonance imaging (fMRI) 664
functional neuroimaging 662–5
functionalism 626, 722, 726, 793

G

Gabler, Siegfried 516
Gaelic 249, 250, 263, 445, 454, 474, 528, 760–1
Gaeltacht 528
Gafaranga, Joseph 140, 142, 400
Gal, Susan 78, 297, 298, 300, 301, 302, 303, 441, 445, 446, 540, 566, 567, 613, 674, 676, 677–8, 683, 688
Galician 394, 422, 528
Gallois, Cindy 159, 163
Gambier, Yves 790, 791
games, language 126
gaming 420, 756
García, Ofelia 140, 175, 267, 309, 322, 337, 402, 406, 413, 414, 686, 719, 723, 726, 727
Gardani, Francesco 15, 17, 64, 65, 68
Garde, Paul 274
Gardner, Robert C. 236
Gardner-Chloros, Penelope 646
Garifuna 447
Garrett, Andrew 66
Garrett, Paul 668
Garrett, Peter 223, 236, 237, 239, 434
Gassner, Denise 138
gating techniques 658
Gawrisch, Vanessa 550
gaze 140
Gazzola, Michele 708–9, 710, 713, 716
Gbe 27, 30, 103
Gbogi, Michael Tosin 325–6
Geeraerts, Dirk 42, 53, 54, 60, 385
Geertz, Clifford 590, 613, 615
gender (people's) 246–56
– code-switching 164
– and differences in learning languages 194
– and ethnicity 247, 250, 252, 286
– identity 686
– language policies 362
– and language shift 446–7
– social network analysis (SNA) 565
– and social networks 446
– social psychology 779
– and speech communities 313
– urban-rural contact 327
gender agreement (grammar) 28, 31, 55, 70, 198, 200, 251, 542
generalizations 154, 187, 188, 445, 487, 509, 643
generative linguistics 30, 200
Genesee, Fred 191, 407, 408, 409
genetic data 14, 20, 72
Genette, Gerard 749
genocide 462, 737
geography
– geographical boundaries 311
– geographical diffusion 118
– geographical identity and ethnicity 286
– urban and rural definitions 320
– urban-rural contact 321
geolinguistics 423, 523–35, 551–2, 760
geopolitical borders 262, 270, 273
geosemiotics 140, 435
German
– in Austria 566
– Baltic German 538
– in Belgium 349, 444
– borrowing 53, 60
– borrowing from 149
– as colonial language 700

– in Eastern Europe 538
– Germany 392
– hybrid versions 543
– international language 399
– in Italy 228
– Kroondal German 459
– in Luxembourg 265
– as a minority language 537
– in Namibia 458–9, 701
– orthography 78, 299
– pragmatics 138
– script choice 79
– in Switzerland 349, 398, 537, 552, 628–9, 712
– varieties 113, 114, 115
Germany 364, 392
Gervain, Judith 189
gesture 139, 140
Ghanian Pidgin English 70
Gibbons, John 164
Giddens, Anthony 332
Giezen, Marcel R. 217
Gilbert, G. Nigel 629
Giles, Howard 117, 159, 163, 178, 179, 203, 235, 236, 384, 444, 456, 553, 779, 780, 781
Gilmore, Perry 591
Ginsburgh, Victor 707, 709
Gitlin, Todd 755
Glaser, Barney 494
Glaser, Clive 325
Gleitman, Leila 38
Glick-Schiller, Naomi 687
globalization
– domains 541
– endangered languages 550
– ethnography 592, 593
– global capitals 772
– and homogenization 273
– language boundaries 273
– language economics 715
– language ideology 301, 303, 304
– language policy and planning (LPP) 362
– legal studies 738
– linguistic landscapes 433
– and linguistic repertoires 287
– media 754–5
– minority language rights 371, 381
– mixed languages 324
– multilingualism 258, 266
– multi-site ethnography 619
– national language ideologies 352
– nexus analysis 583
– societal aspects of language contact 7
– sociodigital space 754, 759
– Spanish 303
– standardization 386
– superdiversity 165, 332–43
– translation 750
– urban-rural contact 320
glossophagia 769
Glynn, Kevin 424
Gobbo, Federico 128
Gode, Alexander 127
Goebl, Hans 1, 2
Goethe-Institut 361
Goffman, Erving 138, 141, 164, 279, 576, 615–16, 620, 633
Gogolin, Ingrid 165, 364
Gold, Ann Grodzins 311
Goldvarb 482
Gollan, Tamar H. 213, 660
Gombert, Jean-Emile 224
Gómez de Garcia, Jule 446
Gómez Rendón, Jorge 66
Gonzàlez, Isaac 583
González, Norma 589
Gonzo, Susan 204
Good, Jeff 311
Gordon, Matthew 514, 518, 563, 565
Gorter, Durk 240, 360, 375, 407, 434, 436, 555, 594
Government Constraint 159, 160
Graber, Kathryn E. 687
Graded Intergenerational Disruption Scale (GIDS) 444, 456, 555
Graffman, Katarina 760–1
Grainger, Jonathan 211, 213, 658
grammar *see also* morphology
– age of speakers 172
– bilingual child language acquisition 190
– borrowing 114, 153
– code-switching 160–1, 162, 163–4
– constructed languages 126–7, 129
– contact-induced change 14
– creoles 102, 107
– and language attrition 200
– pidgins and creoles 101, 106
– process-centered approaches 198
– standardization 388–9
– syntactic change 24–38
grammaticality judgment tests 225, 655–6, 663
grammaticalization 40, 43–4, 102, 115, 150

Granovetter, Mark S. 562, 565
Grant, Anthony P. 2, 53–4
graphemics 76–88
graphics 435, 471–8, 594–5
Grassfields, Cameroon 276–7
Greek 53, 55–60, 65, 149, 150, 154, 261, 344, 505, 537, 646
Greek Australian English 287
Greenacre, Michael 475, 484
Greenfield, Lawrence 399
Greenwald, Anthony G. 238
Grenoble, Lenore 66, 446, 550, 555, 674, 682
Grice, Paul 504
Gricean maxims 163
Grimshaw, Jane 32
Grin, François 6, 708, 710, 712, 713, 715, 716, 734, 736, 738
Grondelaers, Stefan 60
Grosjean, François 166, 177, 202, 259, 406, 658, 719
Grossman, Eitan 18, 19
grounded theory 494, 495, 592
ground-up studies 590
group interviews 606–7
group second language acquisition 177
group status 783–4
Groves, Robert M. 520
Guaraní 474–5, 476–8, 538, 549, 644, 698
Guardado, Martin 603
Guernesiais 568
Guitiérez, Xavier 223, 224
Gujerati 459–61
Gullberg, Marianne 643
Gumperz, John 47, 66, 67, 138, 159, 163, 164, 308, 309, 311, 312, 323, 327, 339, 402, 441, 445, 488–9, 493, 540, 564, 566, 589, 602, 606, 615, 683, 684
Gungbe 27
Gurindji 19
Gurindji Kriol 19, 68, 644
Gutiérrez-Rexach, Javier 28
Guttman, Louis 478–9
Guyana 310
Guzmán, Gualberto A. 645

H

Haarmann, Harald 360, 553
Haberstroh, Susanne 505
habituation-dishabituation paradigm 654
habitus 313
Haboud, Marleen 592–3
Hachimi, Atiqa 327
Hainanese 263
Haitian 70
Haitian Creole 106
Halfacree, Keith 320
Hall, Robert A. 100, 685
Halverson, Sandra L. 798
Hambruch, Paul 697, 698
Hamel, Rainer Enrique 402
Hamers, Josiane 226
Hamid, M. Obaidul 239
Hammarström, Harald 14
Hancock, Ian F. 54
Hanks, Peter 60–1
Hansen, Bjórn 115
Harlem, New York 446, 564
Harré, Rom 629
Harris, Alice C. 25, 27
Harris, Roy 543
Harvey, Graham 618
Haspelmath, Martin 27, 42, 54, 57, 83, 152
Haugen, Einar 1, 2, 25, 53, 64, 88, 90, 139, 148, 149, 198, 199, 322, 346, 358, 359, 386, 391, 441, 444, 542, 550, 553, 588, 593, 719, 720, 737
Hausa 442
Hawaii 758
Hawaiian Creole 102, 457
Hawkins, Eric 223
Hayashi's quantification 475
Haynie, Hannah 14
Hayward, Richard 39
Heath, Jeffrey G. 56, 64, 65
Heath, Shirley Brice 589
heavy borrowing 152
Hebrew 80, 125, 129, 150, 153, 154, 347, 432–3, 538, 793–4
hedges 251
Heil, Anett 131
Heine, Bernd 40, 43, 44, 66, 150, 172, 462, 463
Heinich, Patrick 455
Heller, Monica 301, 406, 445, 490, 634, 683, 719
Hélot, Christine 412
hemodynamic magnetic signals 664
Henne-Ochoa, Richard 683
Herasimchuk, Eleanor 468
Herder, Johann Gottfried von 346, 769
Herdina, Philip 222

heritage languages
– age of speakers 177
– anthropology 689
– attrition 198, 201
– in censuses 527
– diglossia and bilingualism 538
– domain analysis 540
– language maintenance 455, 457, 458–61
– language shift 442
– media 423
– methods for studying 540
– morphological change 69
– social boundaries 275
– syntactic change 33
Hermes, Mary 302
Herold, Ruth 115
Herring, John R. 643
heteroglossia 240, 339, 413, 424, 721, 747, 750, 758, 760
heteronymy 392
heterosemy 39
Hickey, Raymond 2, 113, 115
hierarchies, language 300–1, 688
hierarchies of borrowing 15–16, 18, 24–5, 42, 57, 63, 64, 90, 152–3, 155
Hill, Jane H. 445
Hindi 64, 351, 381, 442, 459, 461, 644, 646
Hinskens, Frans 162
Hinton, Leanne 682
historical body 577, 583
historical linguistics 100, 162, 173, 556, 564
Hittite 66
Hjelmslev, Louis 69
Hlavac, Jim 141
Hobsbawm, Eric 347, 371
Hoffman, Michol F. 290, 605, 608–9
Hogan-Brun, Gabrielle 732, 737
Højrup, Thomas 565
Hokenson, Jan Walsh 744
Hokkien 263
Holberg, Ludwig 389
Holistic Ecolinguistic Model for the Analysis of Language Policy (HEMALP) 364–6
Holm, John 104
Holman, Eric W. 14, 20
Holmes, James S. 789, 790
Holmes, Janet 247, 248, 251
Holmquist, Jonathan 605, 609
Holocaust survivors 203
Holquist, Michael 745–6
homogeneity analysis 475
homogenization 118, 300, 301, 335, 371, 384–5, 406, 592, 715, 769
homogenous language, imagined 298–9
homographs 211, 212, 214, 216, 657
homophones 211, 212, 213, 214, 217, 426, 657
Honeybone, Patrick 117
Hong Kong 411
hooks, bell 617–18
Hopi 591
Hopper, Paul 40
Horn, Laurence 136
Hornberger, Nancy H. 357, 359, 407, 589, 590, 593, 595
Horner, Kristine 301, 337
Horvath, Barbara 248, 288
Howell, Robert B. 89, 90
Hroch, Miroslav 346, 768
Hrycyna, Melania 472, 482
Huang, Guowen 116
Hudson, Alan 539
Huebner, Thom 434, 435
Hulsen, Madeleine 568
Hult, Francis M. 359, 577, 582, 593, 594
human rights 362–3, 371, 377, 378, 631, 732, 734, 735–6
humanities, linguistic turn in 9
Humbert, Philippe N. 526, 527, 529
Hungarian 58, 115, 348, 445, 446
Hungary 238, 344, 566
hunter-gatherers 462
Hup 43, 44, 46
Hurrian 66
hybrid social identities 34, 593, 687, 779
hybrid styles 178
hybridization 419, 756, 760, 761
hygiene, linguistic 420, 421
Hyltenstam, Kenneth 454, 456, 555
Hymes, Dell 100, 101, 104, 140, 307, 308, 309, 583, 589, 592, 596, 607, 615, 689
Hyp 588
hyperforeignism 55
hypermasculinity 325, 327
hyponyms 59
hypothesis testing 481–4

I

Iannàccaro, Gabriele 444, 528, 529, 530, 531, 532, 550
Icelandic 56–7, 388

iconicity 82
iconization 78, 297
ideal speaker-listener 258
identification tasks (psycholinguistics) 658
identity 283–96
– adolescents 176
– allegiance to L1 identity 290
– anthropology 685–7
– code-switching 138, 164, 540
– collective/group 139, 767–8, 780–1, 783
– complex diversity 771
– ethnicity 283–4, 314, 686
– fieldworker profiles 608–9
– in-group/out-group 235, 283, 323, 324, 347, 357, 423, 496, 609, 684, 685, 779–80, 782–3
– heritage languages 458, 460
– hybrid social identities 34, 593, 687, 779
– hyphenated identities 785
– iconization 78
– identity politics 542
– ideology 300
– indexed by language 235, 240, 290, 314, 325, 677, 683, 684, 685–7, 737, 757, 768, 769, 774
– indigenous languages 462
– intersectionality 286
– language awareness 412
– language economics 711
– language policy and planning (LPP) 366
– linguistic boundaries 276
– minority language rights 374
– mixed languages 278, 684
– multiple 686
– multivalency of 285
– national 248, 274, 375, 686, 722, 736, 767–8
– negotiation of 686
– and orthography 78–9, 80, 82–3, 85
– place identity 286, 311–12
– postmodernism 313
– pragmatics 138, 139
– researchers' 618–19
– as social construct 685
– and speech communities 311
– transnational 687
– urban-rural contact 327
– and written language 76
ideologies 296–307
– anthropology 687–9
– attitudes 239
– in census data 444
– code-switching 539
– colonialism 697–8, 700, 703
– competing 689
– definition of 297
– discourse analysis 633
– domain analysis 542, 546
– education 411
– emergence as trend 7
– ethnicity 287
– ethnography 590, 593
– gender 246, 247, 248
– geolinguistics 524
– globalization 337
– interviews (research method) 491
– language policy and planning (LPP) 359, 360, 362
– language shift 445, 446–7
– languages as 165
– and the media 422, 424
– metrolingualism 543
– mixed languages 684
– monolingualism 258, 262, 309, 310, 315, 445, 688, 721, 722, 724, 738, 747, 773
– mother tongue 398
– national ideologies 56–7, 343, 347–8, 737
– one nation-one language 261–2, 300, 419
– and power dynamics 235, 298, 300, 302, 309, 688
– purism 280, 299, 301, 349, 362, 364, 389, 420, 421–4, 457, 698, 760
– and research ethics 674–5
– and revival 346
– semantic borrowing 45–6
– and speech communities 309–10
– standard language 297, 299, 302, 390, 537, 609, 688, 722, 755, 756, 768
– standardization 385, 386–7, 388
– and survey methods 529–30
ideology brokers 299
Idi 260
idiolects 117, 120
idiomatic practices 324
Idsardi, William 91, 93–4
Ifukor, Presley 138, 141
Igbo 350–1
Illyrian Movement 347
imageability 60
imagined communities 311, 445, 731
Imbabura Quichua 66
imitation 13
immersion, ethnographic 591, 602

immersion programs 225, 406, 407, 408–9, 410, 724
imperfect learning 18–19, 177
imperialism, linguistic 350, 541, 550, 738
implicational scales 478–80
implicational universals 15
implicatures 43
Implicit Association Test (IAT) 238
implicit learning 224, 225, 240
importations, phonological 88
imposition 13, 25, 26, 90, 92, 93–6, 97, 115–16
imputation, data 521
incipient bilingualism 257
incommensurability 39
Indefrey, Peter 664
indexical meaning 494, 496
indexicality, theory of 247
indexicality of languages 309
India 47, 78, 225, 320, 351, 381, 435, 564, 689
Indian English 447, 484
indigenization, linguistic 350–1
indigenous languages
– anthropology 682
– code-switching 163
– colonial studies 702
– education 721, 722
– ethnography 590, 591
– globalization 324
– international law 734
– language maintenance 455, 457, 461–3
– language politics 363
– language rights and policies 373–4, 379–80
– language shift 447
– and the media 756, 757–8
– multilingual contexts 260, 263
– nation states 350, 351
– nexus analysis 582
– typology 552
indigenous research methods 674
indirect indexicality 251, 252
indirect research methods 236, 238, 240
individual
– communities of practice 314
– epistemological shift towards 314
– factors in language attrition 202
– focus on the 173, 180, 198, 204, 314, 582, 719
– as holder of linguistic rights 736
– individual differences 213
– language policies 361
– and language shift/maintenance 455
– rights of the 377
– translator (versus translation) studies 798
– variation in bilingual lexical processing 210–22
Indonesia 260, 442, 446, 683, 688
Indonesian 442
inequality/injustice 9, 298, 325, 455, 592, 595, 673, 688, 716, 737
infant perception studies 174, 186, 654
inferencing 493
inflection 55, 64, 68, 70, 71, 101, 106–7, 149, 150, 154, 156, 200
inflectional affixes 17, 63, 64–5, 150–2
informative sampling 514
informed consent 491, 607, 669, 675, 676
inheritance 14
inhibitory skills 202, 215, 217
inkhorn controversy 58
innate language facility 165
Inner Mongolia 312
innovation 101, 102, 175, 177, 310 *see also* creativity
Inoue, Fumio 476
input, quality of 194, 225, 569
insertional code-switching 67–8, 116, 141, 161, 162, 163, 190, 338, 646
insider's view 164, 488, 589, 594, 603, 604, 619
institutional contexts 399, 630–1, 633, 737–8, 759–61, 762, 798
institutional language support 203
instrumentalism 236
integration 58–9, 65, 154, 162, 200
integrational linguistics 543
intensity of contact continuum 18, 24–5, 54, 64, 97, 155, 202, 684, 697
interaction order 576–7
interactional analysis 174
interactional cues 336
interactionist research approaches 488–9, 493–4, 539
Interactive Activation framework 211
interactive alignment 166
interactivity 210–22
intercultural communication 4, 42, 136, 141, 249
intercultural pragmatics 137
interdialect forms 116–18, 322
interdisciplinarity
– education 407

– future directions 10
– language economics 707, 709, 714
– language ideology 300, 304
– language politics 362
– linguistic landscapes 431
– multilingual contexts 267
– nexus analysis 574, 579
– politics 766
– pragmatics 136
– qualitative research 488
– social network analysis (SNA) 562
– translation studies 791
– typology 553
interdiscursivity 690
interethnic communication 104–5
Interface Hypothesis 29
interfaces of the language faculty 29
interference
– adolescents 177
– alternatives to traditional education 726
– asymmetrical contact-induced change 13, 18
– children 175
– code-switching 159, 160, 165
– colonialism 700
– education 720, 724
– and language attrition 200, 202
– language planning 364
– phonological change 88
– semilingualism 722
– syntactic change 33, 40, 45
– and translation 790, 793
– varieties in contact 115, 120
interference tasks 662
intergenerational transmission 186, 456, 541, 545
intergroup theories 778–85
interlanguage 30, 31, 97, 136, 137–8, 165, 720
Interlingua 126–7, 128
interlingual boundaries 159, 724, 794
interlingual contagion 322
interlingual effects 798
interlingual identification 726
interlinguistics 124
intermediate creoles 26, 104, 105
intermediate learners 185
international adoptees 191
International Center for Research on Bilingualism 3
international communication 236, 238, 266, 335, 361–2, 399
International Corpus of Learner English (ICLE) 641
International Covenant on Civil and Political Rights (ICCPR) 371, 377, 378–9, 734
International Journal of Multilingualism 434
International Journal of the Sociology of Language 77
international language policies 361–2
international law 371, 372, 373, 376–80, 734–9
internationalist policies 389
internationalization 362
internet-based media *see also* social media
– anthropology 690
– corpora 639–40
– as data source 545
– dominance of English 266
– emergence as trend 8
– language maintenance 457–8
– media linguistics 419, 425–6
– media/communication studies 754–65
– multilingual communities of practice 265
– pragmatics 138
– as source of participants 610
– spread of language 338
– surveys 501, 507, 518
– translation 797
interpretative repertoires (IR) 628–9
interpreting 661
interpreting studies 789, 796
intersectionality 286
intertextuality 631, 633, 690, 795
interviews (research method) 491, 578, 591–2, 600–13, 628–9
intonation 116, 246, 323, 412, 642, 778
intuition 225
invented languages 125
Ionin, Tania 30
Iquito 46
Ireland 528, 563, 564–5
Irish 115, 239, 264–5, 401, 423–4, 540, 544, 556, 713, 758, 760
Irish English 115, 447
irony 240
Irvine, Judith 78, 165, 297, 298, 302, 674, 677–8
Iscamtho/Isicamtho 277, 543
isoglosses 321
isomorphism 47
Italian 56, 58, 78, 83, 84, 228, 550, 712
Italy 446, 527, 528, 551

Iverson, Gregory 91, 93, 94
Izadi, Dariush 582

J

Jackendoff, Ray 29
Jackson, Tony E. 748
Jacquemet, Marco 337
Jaffe, Alexandra 77, 299, 300, 686
Jake, Janice L. 166
Jamaican Creole 106, 479
James, Allen 435
James, Carl 229
James Bay Cree 551
Janda, Richard D. 55
Jansen, Wim 127, 131
Japan 337, 422, 432, 434, 457
Japanese 53, 55, 81, 93, 106, 113, 114, 337, 352
Jaspers, Jürgen 139, 489, 494, 539
Jaworski, Adam 297, 434
Jejueo revitalization movement 455
Jernudd, Björn 358
Jerusalem 432–3
Jespersen, Otto 69
Jessner, Ulrike 222, 223, 226, 227, 228, 229, 230, 257, 267
Jewish Americans 285, 291
Jie, Dong 588, 592
Johanson, Lars 13, 58, 59
Johns, Michael A. 645
Johnson, Daniel E. 484
Johnson, David Cassels 359, 361, 595
Johnson, Gary 606
Johnson, Ian 760
Johnson, Jacqueline 656
Johnson, Sally 299
Jones, Elin Haf Gruffydd 421, 422, 424
Jones, Mari 550
Jones, Rodney H. 574, 580, 582
Jordanian Domari 155
Jørgensen, J. Normann 140, 324, 337, 338, 402
Joseph, John E. 297
Joseph, Michael 593
Jourdan, Christine 102
Journal of Language Contact (JLC) 2
Journal of Pragmatics 140, 141–2
Journal of Sociolinguistics 565
Joyce, James 750
Judt, Tony 274
justice, linguistic 544, 716, 732, 738–9

K

Kachru, Braj 447
Kafka, Franz 743, 748, 749
Kahn, Hana Muzika 605
Kaier, Anders N. 515
Kamtok 104
Kamusella, Tomasz 348, 349
Kamwangamalu, Nkonko M. 325, 326, 327
Kang, Yoonjung 95, 96
Kannada 66, 655, 684
Kaplan, Robert B. 364, 366
Kaqchikel 605
Karmiloff-Smith, Annette 224
Kasaipwalova, 618
Kasanga, Luanga A. 435
Kaska 689
Kâte 697, 698
katharevousa 56
Kaufman, Terrence 13, 15, 18, 24–5, 54, 64, 66, 100, 112, 114, 115, 148, 152, 155, 162, 175, 684, 697, 701
Kazakh 555
Keating, Patricia A. 91, 92
Kecskes, Istvan 136, 227
Keesing, Roger 102
Kellman, Steven 747
Kelly-Holmes, Helen 239, 275, 422, 423, 424, 426, 758, 760
Kelman, James 748–9, 750–1
Kemp, Charlotte 226
Kempe, Vera 186
Kenya 70, 178, 322, 325, 326, 401, 460, 492
Kenyan Pidgin Swahili 70
Kepsu, Kaisa 759
Kerrigan, John 748
Kerswill, Paul 113, 116, 118, 120, 172, 175, 178, 310, 320
Keyser, Samuel Jay 92
Khan, Sadiq 264
Khider, Abbas 749
Khuzistani Arabic 151, 153
Kihm, Alain 70
Kikóngo 70, 71
King, Kendall A. 302, 555, 592–3
King, Ruth 26, 66
kinship terms 57, 153, 251
Kioko, Eric M. 326

Kipp, Sandra 326
Kitúba 70, 71
Klingon 129
Kloferová, Stanislava 117
Kloss, Heinz 1, 347, 359, 391–2, 441, 443, 444, 456, 542, 553
Klöter, Henning 83, 84
Knooihuizen, Remco 556, 557
knowledge-based societies 715, 774
Kochenov, Dimmitry 738
koineization 116, 117, 120, 162, 278, 322, 696
Kolker, Boris 129
Koller, Werner 789, 790
Kong, Shuyu 750
König, Christa 461, 462, 463
Kontaktlinguistik / Contact Linguistics / Linguistique de Contact (Goebl et al. 1996, 1997) 1, 3–4, 5
Kontra, Miklos 732, 737, 738
Koostra, Gerrit Jan 166
Köpke, Barbara 199, 202, 204
Koptjevskaja-Tamm, Maria 39, 43, 46
Korea 336–7
Korean 81, 93, 113, 423
Korlai 70
Koro 45
Kossman, Maarten 17
Kouwenberg, Silvia 100, 107
Kramsch, Claire 173, 582, 583
Kranich, Svenja 790, 792
Kraus, Peter A. 768, 769, 771, 772, 773, 775
Krauss, Michael 443, 445, 550, 670
Kress, Gunter 139
Kri 46
Krio 106, 630
Kristen, Cornelia 193, 194
Kroch, Anthony 180
Kroondal German 459
Kroskrity, Paul V. 297, 300, 309, 447, 674
Kruger, Haidee 798
Kruskal, William 516
Kulick, Don 445
Kuppens, An H. 756
Kupwar Kannada 66
Kupwar village, India 47
Kurdish 149, 150, 154
Kusters, Wouter 18
Kuteeva, Maria 541
Kuteva, Tania 40, 43, 44, 66, 150, 172
Kuzborska, Elžbieta 371, 376, 378, 381
Kven 394
Kymlicka, Will 359, 374, 375, 377, 737, 738

L

L1 acquisition *see* first language acquisition
L1-regression 179–80 *see also* attrition
labor markets 707–18
Labov, William 5, 113, 165, 173, 178, 235, 236, 239, 246, 249, 250, 252, 285, 286, 308, 309, 310, 321, 514, 518, 520, 556, 564, 580, 601, 602, 603, 604, 605, 606, 669, 670, 671, 672
Lacorne, Denis 274
Ladd, Heather 758
Ladin 228, 394, 528, 529, 550
Lado, Robert 137, 139, 165
Lafaye, Beverley Elsom 422
Lago, Sol 656
LAILA 229
LAILA-BICS 229
Laitin, David 323, 350
Lakoff, George 39
Lakon 45
Lambert, Wallace E. 236, 237
Landry, Roger 431, 432, 436
landscapes, linguistic 240, 360, 431–41, 575, 591, 594, 690
Lane, Pia 394, 575, 577, 578, 579, 580, 582, 583
Lang, Marvel 321, 322
language, a, defining *see* named languages, defining
Language Loss conference 199
Language Management Theory 358–9
language policy and planning (LPP) 357–70
Language Variation Suite 484
language-dependent recall 505
languages of wider communication (LWCs) 442–3
languaging 6, 9, 140, 324, 436 *see also* translanguaging
langue 721, 743, 744, 745
Lanza, Elizabeth 174, 431, 435, 436, 568
Lao 41, 46
Lapkin, Sharon 784
Larson-Hall, Jenifer 471
Lasagabaster, David 409, 410
Latgalian 364, 528
Latin 17, 55, 57, 60, 79, 113, 150, 261, 344, 348, 387, 388, 391, 797
Latina girl gangs 602, 686
Latine sine flexione 124, 128
Latour, Bruno 798
Latvia 364

Latvian 362, 538
law and language 540, 731–42
Lawson-Sako, Sarah 164
Le Page, Robert B. 162, 279, 310, 314, 386, 441, 445, 468, 481, 684
leaky diglossia 538–9
learner corpora 641
learner varieties 698, 700
learning strategies 226, 230
Lee, Jamie Shinhee 423
Leeman, Jennifer 434–5, 525, 526, 527
Lefebvre, Claire 33, 71, 102, 103
legal studies 731–42
legitimate language 299
Léglise, Isabelle 311, 701
Lehmann, Winfred P. 57
Lehner, Sabine 633
Lehtola, Veli-Pekka 756
Lemerig 45
Leopold, Werner 187
Leow, Ronald P. 225
Lëtzebuergesch 556
Lev-Ari, Shiri 564
levelling 116, 117–18, 162, 322, 696
Levon, Erez 238
Lewis-Kraus, Gideon 616
lexical decision tasks 657–8, 659, 661
lexical items *see also* vocabulary
– asymmetrical borrowing of 16
– bilingual first language acquisition (BFLA) 191
– contact-induced semantic change 38, 41
– impact of colonialism 698–9, 700
– implicational hierarchy of borrowing 64–5
– individual variation in bilingual lexical processing 210–22
– and language attrition 200 lexical competition 39
– modernization policies 362
– orthographic assimilation of loanwords 83–4
– pidgins 101
– word formation and vocabulary expansion 52–63
lexical layers 54
L-G mixed languages 68
Li, David C. C. 411
Li, Peggy 38
Li Wei 140–1, 174, 175, 189, 267, 319, 402, 410, 413, 446, 489, 490, 566, 567, 568, 569, 726, 727
liberalism 362
Lieber, Michael 347
Lieberson, Stanley 503, 526, 528
lifelong learning 230
life-mode 565
lifespan, contact across the 172–84, 225, 267
lifestyle diglossia 540
Light Warlpiri 68–9
Liljegren, Henrik 43
Lim, Lisa 198, 304, 588
Lin, Angel M. Y 411, 588
Linck, Jared A. 215
Lindstedt, Jouko 130
Lingua Franca (Pidgin Italian) 70
lingua francas
– domains 401, 549
– economics of language 715
– English 238
– globalization 324, 326, 335
– Latin 129
– legal studies 737
– multilingual contexts 260, 261, 263, 266, 267
– politics 766, 771, 775
– Urdu 351
– workplaces 337
lingualism (*Sprachigkeit*) 745–6, 747, 748
linguality (*Sprachlichkeit*) 745–6, 748
linguistic capital 239, 402, 633, 709–10, 713, 715
Linguistic Category Model 782
linguistic communities 309, 312
Linguistic Data Consortium 639
linguistic landscapes (LL) 431–41
linguistic markets 239
Lippert, Sandra L. 79–80
List, Johann-Mattis 48, 72
listedness 162
literacy
– domain analysis 537
– education 407, 722, 723
– metalinguistic awareness 228
– nation states 343, 344, 345–6, 350, 352
– nexus analysis 582
– orthography and graphemics 79
– as a social practice 77
– uninstructed language acquisition 192
literary studies 240, 743–53, 796, 797
Lithuanian 362
Little, Roderick J. A. 521
Liu, Guo-Qiang 249
Liverpool 117, 594
Lo Bianco, Joseph 359, 360

loan translations *see* calquing
loanwords
– asymmetrical borrowing of 16, 17
– colonialism 699
– corpus planning 363–4
– loan blends 53
– and the media 423
– orthographic assimilation 83–4
– phonology 88, 90
– semantic borrowing 42
– vocabulary expansion 53
local imposition 94–6
LOCALE 539–40
localization 421
Logical Form (LF) 31
logistical regression analysis 483
Lohr, Sharon L. 517, 519, 521
Lojban 124
London 164, 175, 177, 251, 263, 264, 266–7, 289, 460, 493, 684, 772
Long, Michael H. 45, 174
longitudinal studies 7, 20, 137–8, 187, 192, 193, 194, 445, 503, 541, 589
Longobardi, Giuseppe 30
López, Alexis A. 414
Lorenzo-Dus, Nuria 137
loss, domain 399–400, 536, 541
loss, language 372, 737 *see also* attrition; death, language
Lou, Jackie Jia 140, 584
Lourie, Margaret A. 444
Loveday, Leo J. 310
Low and High varieties 163, 263, 537–9, 770 *see also* diglossia
Lowland Mayan languages 43
LPC effect 217
Lucy, John A. 685
Lüdi, Georges 140
Luís, Ana 65
Lüpke, Friederike 311, 442, 698
Luxembourg 265
Luxembourgish 265, 301, 426, 528
Lynch, Andrew 250

M

Ma, Roxana 468
Ma'a 684
Maasai 461–2
Maass, Anne 782, 783
Macaro, Ernesto 406
Macedonia 298, 348
Mackey, William F. 1, 3, 401, 443, 723
MacPherson, Seonaigh 249
MacSwan, Jeff 159, 161, 192, 643
Madsen, Lian Malai 139, 402, 491, 495
Maguire, James 461
Mahootian, Shahrzad 161
maintenance of languages 454–67
– anthropology 683, 689
– attrition 202, 204
– code-switching 163
– colonial studies 697
– domain analysis 540, 541
– gender 248
– language politics 362
– language shift 444, 447
– legal regulations 732
– media 756
– and the media 757–8
– social network analysis (SNA) 565, 568
– syntactic change 26
– urban-rural contact 324
Makoni, Sinfree 173, 174, 177, 179, 180, 687, 726, 748
Malamatidou, Sofia 791
Malay 69, 262, 263, 540
Malayalam 263
Malaysia 262, 351, 540
Malinowski, Bronisław 489–90, 588–9, 595, 614–15, 616, 617, 618
Maltese 58, 82, 153, 556
Maltese English 115
Maly, Ico 437
Mańczak, Witold 42
Mandarin 263, 301, 423, 460, 594
Mandelbrot, Benoît 272, 273
Manley, Marilyn S. 479
Mano, Winston 421, 422
Māori 363, 373, 424, 528, 606, 758
map tasks 239
Maphosa, France 323
maps, geolinguistic 531–2
Marathi 66, 684
March, Lucien 515
Marian, Viorica 211, 216, 217, 657
Maritime Pidgins 105
markedness 39, 42, 56, 448
Markedness Model 159, 163
Markelin, Lia 757
market values 711
marriage, exogamic 19, 34, 260, 277, 459, 460, 683

Marten, Heiko F. 364, 365, 436
Martha's Vineyard 239, 286, 601
Martin, Peter 410
Martín-Beltrán, Melinda 413
Martinet, André 126
Martin-Jones, Marilyn 406, 410, 582, 723
Marxism 239, 716
Maryns, Katrijn 629–30
masked phonological priming 658
Massive Multiplayer Online Games 420
Mass-Observation Project 617
Mastumoto, Kazuko 568
Matched-Guise Technique (MGT) 237
matched-guise techniques 164, 238, 310, 446
Matisoff, James 39, 40
MATLAB 484
Matras, Yaron 1, 2, 13, 15, 17, 19–20, 40, 45, 60, 64, 68, 77, 89, 93, 97, 114, 116, 148, 149, 150, 151, 152, 153, 155, 172, 198, 278, 334, 335, 336, 677, 684, 766
matrix data representation 470–1, 475
matrix language 67, 68, 148, 154, 161, 166, 277
Matrix Language Frame (MLF) 67, 159, 161, 166, 643
Matrix Language Turnover Hypothesis 162
Matsuda, Paul Kei 747
MAT(ter) borrowing 13, 15–16, 40, 64–6, 68–9
Matthey, Marinette 412
May, Stephen 267, 321, 371, 372, 376, 377, 378, 379, 381, 407, 544, 733, 738, 739, 769
Maya 299
Mayan 43, 426, 685
Mazzucto, Valentina 565
Mbugu 684
McCarty, Teresa L. 588, 593
McConvell, Patrick 19, 20
McCormick, Kay 164, 435
McEntegart, Damian 481
McKenzie, Robert M. 237, 238
McLaughlin, Barry 226, 266, 442
McLaughlin, Fiona 618, 620
McLuhan, Marshall 418
McMahon, April 56, 57
McRae, Kenneth D. 359
McWhorter, John 26, 27, 70, 71, 100, 105, 106, 176, 685
Meakins, Felicity 19, 20, 65, 68, 69, 72, 278
meaning *see also* semantics
– connotational meaning 60, 114
– contact-induced semantic change 38–52
– diachrony of meaning 39–40, 48
– encyclopedic views of meaning 39
– entailed meanings 39, 40, 43
– form-meaning relationship 38, 40–3
– indexical meaning 240, 494, 496
– in interaction 490
– interactionist research approaches 494
– meaning-in-interaction 141
– negotiation of 413
– and orthography 77, 78
– perception of 186
– social meaning 78, 240, 384, 446–7, 683
– sociopragmatics 139
– translation 795
Meänkieli 393–4
Meara, Paul 203
media 418–30
– as cultural texts 333
– and differences in learning languages 194
– heritage languages 458
– indexical meaning 240
– as language gatekeepers 754–5
– and language ideology 304
– language policy and planning (LPP) 361
– as source for learning a foreign language 193
– and standardization 390
– translation 795
Media Lengua 34, 162, 278, 684
media/communication studies 754–65
mediated borrowing 25, 26
mediated discourse analysis (MDA) 574
mediational means 574, 576, 580, 583
medium of instruction 266, 376, 381, 399, 400
Mednyj Aleut 68–9
Meek, Barbra A. 591, 613, 682, 689
Meesters, Gert 424
Meeuwis, Michael 164
Megleno Romanian 18
Meisel, Jürgen 186, 189, 193
Melanesia 568
Melanesian Pidgin 104, 105
melting pots 275
memory 211, 213, 214, 215, 505, 508, 577, 660, 661, 796 *see also* working memory
Menard-Warwick, Julia 249
Mendoza-Denton, Norma 285, 287, 288, 314, 602, 609, 686, 689–90
Menken, Kate 414
Mensah, Eyo 324, 325

Mercier, Julie 216, 217
Mesthrie, Rajend 459, 605
Met, Myriam 409
metacognition 223, 230
metadata 642
metalanguage 297
metalinguistic awareness 56, 222–34, 299, 336, 412, 526, 609
metaphor 42, 60, 173, 324, 335, 628, 794–5
metaphorical code-switching 164, 401, 402, 489, 683
metatypy 47, 69
metonymy 42
metrolingualism 140, 267, 324, 337, 401, 538, 542–3
metro-normativity 321
Mexican Americans 285, 289, 686
Meyerhoff, Miriam 6, 103, 250, 251, 252, 603, 604, 607
M-factor 228–9
Miceli, Luisa 45
Michael, Lev 19, 34
Michaelis, Susanna 70, 108
Michif 34, 68–9, 278, 684
Middle High German 115
Midgley, Katherine J. 213
Migge, Bettina 103, 107, 311, 701
Mignolo, Walter 727
migration *see also* superdiversity
– and age 176, 192
– anthropology 690
– attitudes 238, 239, 352
– attrition 198, 199, 203, 204
– and bilingualism 175
– colonial studies 702–3
– corpus linguistics 640, 644
– and dialects 116, 117–18, 120
– economics 709–10
– education 724–5
– ethnography 592–3
– foreign language learning 192–3
– future directions in research 10
– gender stratification 250
– geolinguistics 525
– globalization 333
– ideology 301
– integration through language discourse 628
– language maintenance 456
– language policy and planning (LPP) 362
– and language shift 442, 447
– and linguistic rights 736 and the media 759
– mixed languages 324
– multilingual cities 264
– and multilingualism 260
– multi-site ethnography 619
– national identity 375
– no chapter on 4
– passing 323
– political science 770
– pre-migration capital 193, 326
– and regional phonological adoption 289–90
– social indexicality 249, 251–2
– social network analysis (SNA) 565–7, 568
– societal aspects of language contact 7
– speech communities 310, 311–12
– translanguaging 413
– translation 794
– transnational communities 687
– urban-rural contact 320, 326
Mi'kmaq 481, 618
Milani, Tommaso M. 426, 436
Milburn, Trudy 312
Miller, D. Garry 53, 57
Milroy, James 54, 299, 310, 321, 445, 565
Milroy, Lesley 116, 299, 310, 311, 314, 321, 402, 445, 489, 514, 518, 563, 564–5, 567
Minimalism (syntax) 6, 161, 643
minority language groups
– anthropology 682, 686
– attitudes 235, 239
– autonomy 773
– diglossia and bilingualism 537–8
– discourse analysis 631
– economic value of languages 710–11, 713
– education 406, 407, 719, 724–5, 726, 784
– elicitation of variables 606
– extrinsic minorities 555
– France 363
– globalization and superdiversity 324, 326
– human rights 631
– identity 286
– ideologies 299, 301
– language policies 360
– language policies and rights 363, 370–84
– language policy and planning (LPP) 362
– language politics 769
– language polls 528
– language shift 442, 447

– language vitality 550
– and the law 731, 732
– linguistic rights 735
– maps, geolinguistic 531–2
– and the media 421, 423
– media/communication studies 754–65
– minorities within minorities 551
– multilingual contexts 262, 263
– nation states 343
– national language ideologies 347–8
– official multilingualism 734
– political science 770–2
– sociology-of-language framework 554
– translation studies 794
– typology 551, 555–6
Minority Protection Scheme 376
Miraña (Boran) 46
Mirandese 394
Mirshahidi, Shahriar 238
missionaries 81, 618, 697, 699, 702
Mithun, Marianne 18
mixed grammars 26–7, 67, 159
mixed languages
– Africa 324–6
– anthropology 684
– borrowing 148, 156
– code-switching 160, 162, 164
– colonial studies 702
– contact-induced change 19
– contact-induced syntactic change 26
– corpus linguistics 645
– domains 401
– hybrid social identities 34
– interviews (research method) 606
– lifespan, contact across the 175
– linguistic boundaries 278
– linguistic landscapes 434
– morphology 68–9
– sample analysis 496
– semantic borrowing 45–6
– social boundaries 277, 278
– social media 426
– third spaces 687
– Turkish-German 543
– urban-rural contact 322
mixed utterances (bilingual children) 190
mixed-effects models 468, 483–4
mixed-methods approaches 634
mixed-mode survey design 509
Mobilian Jargon 105
Modan, Gabrielle 434–5
modeling
– contact-induced change 16, 18–20
– corpus linguistics 643
– dialect/standard constellations 391
– economics of language 710
– language shift 444–5
Modern Language Association 723
Modern Standard Chinese 41
modernism 340, 721–5, 726
modernization policies 721, 723
modes, language 202
modular interfaces 27, 29, 33
modularity 92–7
Moffatt, Suzanne 564
Mohanty, Ajit K. 225
Molis, Michelle 656
Møller, Janus Spindler 542
Mondada, Lorenza 141
monism, linguistic 745 *see also* named languages, defining
Monnier, Daniel 432
Monogenesis Hypothesis 28
monoglossia 406, 413, 688, 721, 724
monolingualism
– adult speakers 180
– 'affordability' of 766
– bilingualism through monolingualism 408–9, 410
– education 406, 411–12, 413, 414, 719
– France 363
– gender 252
– and globalization 339
– ideologies 258, 262, 309, 310, 315, 445, 688, 721, 722, 724, 738, 747, 773
– increasingly rare 310
– lexical borrowing 57
– linguistic equivalence 792
– literary studies 744, 746, 747–50
– London 264
– and the media 419, 760
– metalinguistic awareness 224, 229
– monolingual mode (bilinguals in) 166
– monolingual second language programs 725
– multilingual contexts 258, 262
– official monolingualism 734
– and postmodernism 542
– social boundaries 273, 280
– speech communities 308
monophthongization 115, 289
Montanari, Simona 226
Montenegrin 347, 392

Montrul, Silvina 30, 31
mood paradigms 65
Moore, Robert E. 670
Morales, Luis 657
Moravcsik, Edith A. 15, 64, 155
Moretti, Bruno 550
Morford, Jill 215
Morgan, Marcyliena 309, 310, 312, 314
Moriarty, Mairead 422, 436, 760
Moring, Tom 754, 756, 757, 758, 759
Morocco 327, 609
morphology
– bilingual child language acquisition 190
– borrowability 17–18, 42
– borrowing 40, 54–5, 63–76, 149–50, 153, 156, 646, 699
– bound morphology 15–16, 18, 65, 70, 101, 106, 155, 160
– code-switching 160
– colonialism 699
– constraints on borrowing 15
– contact-induced change 14
– contact-induced semantic change 46–7
– contact-induced syntactic change 29–32
– content versus system morphemes 67
– integration of loans 154
– pidgins and creoles 101, 106
– transfer 63–76
– word formation and vocabulary expansion 53
morphology/semantics interface 27, 31–2, 33–4
morphology/syntax interface 65
Mortensen, Janus 401
Mosiro 461
Mosteller, Frederick 516
'mother-tongue' 105, 262, 301, 323, 346, 398, 526, 529, 722
motivations for language learning 238
motivations for linguistic choices 250, 314, 326, 455, 456, 540
Motu 260
Motwap 45
Mougeon, Raymond 173, 186
Moya Guijarro, Arsenio Jesús 142
mTm: A Translation Journal 794
Mu, Yingying 314
Mufwene, Salikoko 2, 26, 27, 70, 71, 100, 333, 334, 442, 549, 701
Mugford, Gerrard 138
Mühlhaüsler, Peter 556, 696, 699, 701
Mulkay, Michael 629
multicompetence frameworks 140, 226, 228
Multicultural London English 175, 177, 609, 686
multiculturalism 264, 285–6, 375, 437, 542, 749, 771
multidisciplinarity 2, 9–10, 587, 696, 762 *see also* interdisciplinarity
multiethnic communities 260–3, 264, 265, 275, 276, 283–4, 325
multiethnolects 175, 177, 178, 286–7
multilevel analysis 632
Multilingual Student Translation Corpus 641
multilingualism *see also* bilingualism; plurilingualism; repertoires, linguistic; trilingualism
– across the lifespan 172–84
– anthropology 683
– and borrowing 66
– cognitive processes 44–5
– communities of practice 264–6
– continuum of multilingualism 259, 309, 310
– defining 258–9
– and dementia 179
– distribution and incidence 259–66
– domains 398, 400, 401
– economic value of languages 713
– education 406–18, 719, 721, 774
– and empire 260–1, 344–5, 348, 695
– empires 260–1
– and a European identity 352
– fluid multilingualism 336–40
– focus on the individual 173
– fractured multilingualism 687
– and gender 247–8
– geolinguistics 523–35
– intercultural communication 141
– internet 138
– lack of distinct separate languages 542
– language economics 712, 715
– language ideology 301
– language of research 592, 605, 676
– language policy and planning (LPP) 363
– language shift 441–54
– lexical borrowing requires restricted 57
– linguistic boundaries 275–7
– linguistic landscapes 431–41
– linguistic nationalism 346–8
– literary studies 744, 745
– measures of multilingual proficiency 227–8, 414
– media 418–30

– metalinguistic awareness 222–34
– mixed languages 324
– multilingual media teams of production 424–5
– multilingual works teams 541
– and multimodality 139
– networked multilingualism 84, 419–20
– nexus analysis 573–87
– official multilingualism 349, 733–4
– political science 767–75
– pragmatics 137–8
– qualitative research 488–9
– reconstituted multilingual areas 277–8
– and semantic change 43, 44, 45–6
– setting language boundaries 272–8
– small-scale 259–60
– societal aspects of multilingualism 257–70
– and speech communities 309
– states 261–3
– stigmatization of 46, 263, 301, 410, 445, 530, 588, 688
– and superdiversity 165
– translation studies 791
– uninstructed language acquisition 185–97
multiliteracies 4, 77
multi-method analysis 488, 489
multimodality
– anthropology 690
– cross-linguistic influence 215, 216, 217
– ideologies 304
– linguistic landscapes 435
– literary studies 747
– nexus analysis 581–2, 583
– pragmatics 139, 141
– research methods 425
– translation 797
multimorphemic expressions 40–1
multinational corporations 238, 240, 266, 361, 419, 420, 424–5, 541, 737–8
multiple correspondence analysis 476
multi-site ethnography 619–21
multitasking 756
multivariate analysis 444, 468, 471, 472, 478–81, 709
Mumin, Meikal 78
Munro, Jennifer 70
Munson, Marcella 744
mutually intelligible varieties 112–13, 117, 278, 345, 391–2
Muysken, Pieter 20, 25, 27, 34, 45, 60, 64, 67, 68, 103, 159, 160, 161, 162, 163, 198, 684
Myers-Scotton, Carol 13, 19, 34, 60, 67, 68, 138, 148, 153, 154, 159, 161, 162, 163, 166, 198, 204, 323, 540, 643
Myles, Florence 192

N

N|uu 462–3
N400 213, 663
Nagy, Naomi 288
Nahir, Moshe 362
Nahuatl 261, 350, 445
named languages, defining
– in censuses 526
– constructed languages 126
– domains 402
– ideologies 297, 298–300
– intralingual identifications 726
– legal definitions 732
– linguistic landscapes 434, 435
– literary studies 745
– multilingual contexts 259
– nation states 347
– pragmatics 139
– rejection of concept of ‘a’ language 542, 687
– and research ethics 677
– social boundaries 271–2
– speech communities 308
– superdiversity 334
Namibia 458–9, 701
naming tasks 659
Nasa Yuwe 422
Nassenstein, Nico 324, 325
Nation, Robert 226
nation states 343–56
– anthropology 686
– education 406
– globalization 326
– ideologies 56–7, 298, 300
– Language Law 733
– language maintenance 457
– language politics 360–1
– language rights and policies 370–1, 377, 380
– language shift 445
– and languages 140–1
– legal studies 733, 736, 737

– linguistic boundaries 276
– and multilingualism 259, 261–3
– one nation-one language 261–2, 300, 419
– political science 767–8
– speech communities 308, 311
– standardization 386
national identity 248, 274, 375, 686, 722, 736, 767–8
national language bodies 359, 360, 364, 736
national languages 343–56
nationalism 82, 343, 345–8, 357, 362, 363, 371, 372, 389, 457, 731, 746, 747, 768
nation-building 261, 275, 360, 386, 754, 767–8
Native Americans 363, 373, 589, 683
native speaker competence, as goal 229, 258
nativization 29, 42, 54, 55, 56, 58–9, 120
Natural Language Processing 469, 646
natural language processing technologies 641
natural phenomena terms 57
naturalism, sociolinguistic 302–3
naturalistic settings 185, 187, 188, 193, 604, 607, 646
Nauruan 698
Navajo 350, 363
Nayak, Nandini 226
Ndyuka-Trio Pidgin 102
Neal, Zachary 311, 312
Nebrija, Antonio de 260
Nekvapil, Jiři 366
Nelde, Peter H. 1–2, 3, 5, 625, 634
Nen 260
neoliberalism 266, 708
neologisms 53, 153, 349, 389, 420 *see also* coinages
network analysis 402, 518, 544–5 *see also* social network analysis (SNA)
Network Strength Scale 565
networked multilingualism 84, 419–20
neuro-cognitive processes 160, 166, 210–22 *see also* cognitive processes
neuroimaging 662–5
neurolinguistics 5, 204, 210
New Literacy Studies 77
New Mexican Spanish 482, 645
New Persian 65
'new speaker' movement 590
New York 153, 160, 265, 285, 287, 291, 446, 564
New Zealand 528, 568, 606
newborn language abilities 186
Newcastle 118, 564
Newport, Elissa L. 656
nexus analysis 435, 573–87
Neyman, Jerzy 515
Ngarluma 71
n-grams 639
Nguni languages 298
Nguu 462
Nguyen, Trang T. T. 239
Nicholas, Sheilah E. 591
Nichols, Johanna 14
Nicolaï, Robert 2, 271, 279
Nigeria 79, 81, 139, 322, 325, 350–1, 366
Nigerian English 138, 139, 350, 447
Nigerian Pidgin English 104, 139, 322, 324–6
Nivaclé 260
Nkep-Bislama 644
Nominal Mapping Parameter (NMP) 29–30
nominal morphology 30, 55, 56, 70, 151
nominalizers 17
nonce formations 60, 162, 483
nonmarket values 711, 713–15
nonresponse 502, 507, 508, 520–1
nonverbal communication 507, 543
Norman French 17
NORMs (non-mobile, old, rural males) 321
Norms and Exploitations, Theory of 60–1
Norris, Sigrid 574, 578, 580, 582, 583
Northern Athapascan 689
Northern Cities Shift 289
Northern Khmer 556–7
Northern Sámi 374, 393, 757, 758
Nortier, Jacomine 175, 177, 277
Norway 105, 118, 262, 320, 364, 374, 386–91, 393–4, 515, 568, 756, 757
Norwegian 64, 88, 90, 198, 359, 364, 386, 388, 390, 392
Norwich 117
noticing 225
nouns
– classification systems 41
– discourse analysis 633
– nominal morphology 30, 55, 56, 70, 151
– nominalizers 17
– perception of 685
novels 743–53
null hypothesis 475
Null Subject Parameter (NSP) 32
null subjects 32
number agreement 28, 30, 31
numerals 43, 127, 152, 153
Numidiam script 82

Nuremberg Code 668–9
nursing homes 180

O

Ó Riagáin, Pádraig 401, 402, 540, 544, 545
Oberwart 566
O'Brien, Gisela 192
observation (research method) 490–1, 514, 569
observer paradox 507–8, 580, 602, 604, 606–7, 608–9, 615, 621, 669
observers observed 617
obsolescence 447–8, 606, 683
Occitan 552
Ochs, Elinor 247, 251, 252
Odlin, Terence 45
official language polls 525–6, 528
official status languages
– anthropology 684
– domain analysis 393
– ideologies 300
– language politics 363
– language rights and policies 371
– legal studies 732, 736, 738
– linguistic boundaries 273
– media/communication studies 761
– multilingual contexts 261, 262, 263, 265, 266
– nation states 344, 345, 350
– orthography and graphemics 79, 83
– politics 771
Ohala, John J. 92
Ojibwe 70
Okada, Yusuke 141
Oksaar, Els 1
Old English 44, 79
Old Norse 54–5, 387
older people 178–80, 185, 225, 250
Oller, D. Kim 191, 194
one nation-one language 261–2, 300, 419
one parent-one language 361
online communities 312 *see also* social media
online ethnographies 690
online forums 138
onomasiological change 54, 60
Onumajuru, Emeka M. 325
Onumajuru, Virginia C. 325
open data 641, 642, 671–2
Opitz, Cornelia 202
Oppenheim, Abraham N. 235
optimal scaling 475
optimality-theoretical constraints 20
oral translation 796
orders of indexicalities 496
Oromo 262, 435
Oroqen 312
orthography 76–88
– colonialism 699–700
– constructed languages 129
– cross-linguistic influence 211, 212, 215, 216
– diacritics 78, 85, 758
– domains 402
– ethnography 594–5
– ideologies 299
– language mixing 495–6
– and the media 422, 758
– nation states 345, 346, 349
– orthographic reforms 80–2
– social media 426
– standardization 388
– transcription of corpora 642
Ortner, Sherry B. 313
Oskaar, Els 778
Oslo Recommendations regarding the Linguistic Rights of National Minorities 734
Ossa Parra, Marcela 414
Ossetian 549, 551
Otheguy, Ricardo 140, 250, 252, 414, 609, 726, 727
Othering 595, 616, 686, 790
Otomí 644, 699
Otsuji, Emi 140, 324, 337, 338, 401, 490, 491, 538, 542, 543, 575, 582
Ottawa-Hull corpus 471, 604
outsider researchers 609, 616
outsider system morphemes 67–8
Overbeek, Greet 321
overgeneralization 198, 201, 448
overt subject pronouns 32–4, 250
ownership, researchers' 671–2

P

P600 663
Pacific Pidgin (English) 105
pair interviews 607
Pakendorf, Brigitte 20, 65
Pakistan 351

Palau 568
Palenquero 70
Palestinian Domari 152, 153, 154
Palosaari, Naomi 448
Pan, Yuling 580–1
Panjabi/Punjabi 352, 493, 686
Pantos, Andrew J. 238
Papp, Tünde 227
Papua New Guinea 43, 47, 259–60, 445, 542, 615, 697, 698
paradigm borrowing 65
Parafita-Couto, M. 643
Paraguay 409, 474–5, 538, 698
parallel corpora 641
parallel language use 390
Paris 433, 772
Park, Joseph Sung-Yul 335
Parkin, David 401
Parkvall, Mikael 70, 71, 131
parody 240
partial accommodation 118
participant observation 236, 488, 591, 594, 595, 613–25, 690
participant perspectives 490, 496
participant research 579 *see also* ethnography
particle placement 484
Parts of Speech tags 642
Parvaresh, Vahid 138
Pasch, Helma 79, 82
passing 322–3
passive constructions 114
passive networks 566–7
past participles 68
Paston Letters 564
Patrick, Donna 320
Patro, Jasabanta 425
Patten, Alan 359, 737, 738
PAT(tern) borrowing 13, 16, 40, 66, 69, 149
pattern replication 149
Paulston, Christina B. 326, 555
pausing behavior 646
Pauwels, Anne 248
Pavlenko, Aneta 6, 176, 200, 247, 338–9, 583, 687
Peano, Guiseppe 124
Pearson's *r* 474
pedagogy 410–14 *see also* education
Pedersen, Karen Margrethe 401
peer group influences 175–7
Peña Sarmiento, María Fernanda 422
Pennsylvania 115
Pennycook, Alistair 140, 173, 308–9, 310, 311, 324, 336, 337, 338, 401, 490, 491, 538, 542, 543, 575, 582, 687, 726, 748
perception studies 290
perceptual cues 92
perceptual dialectology 239
Pereltsvaig, Asya 131
performativity 620–1, 750, 760
Péronnet, Louise 481
Perrin, Daniel 421, 425
Persian 65, 151, 349
personality principle 735
Peru 27, 479, 592
Peterson, David J. 124
Peti-Stantić, Anita 349
Petré, Peter 180
Pfaff, Carol W. 161
phase theories 346
phi-features 30–2
Philadelphia 118, 289, 290
Philippe, Benoît 130
Philips, Susan U. 589
Phillipson, Robert 735, 738, 739
philology 743–53
philosophical languages 128
philosophy and pragmatics 136
phonetics 88–99
phonology 88–99
– attrition 201
– borrowing 45, 54, 114, 149, 150, 154
– chance phonological similarity 18
– children 175
– code-switching 161–2
– colonial studies 700
– constraints on borrowing 45, 54
– constructed languages 128
– cross-linguistic influence 189
– dialect variations 117, 118
– divergence 47
– elicitation of variables 601, 605
– and ethnicity 288–90
– ethnolects 287
– first to be affected by contact-induced change 18
– gender stratification 250–1
– homogenization 118
– imposition 115
– languages shift 447
– neuro-cognitive processes 211
– orthographic assimilation of loanwords 83–4
– pidgins and creoles 101, 107

– semantic borrowing 40
– shift 448
– social indexicality 238
– social network analysis (SNA) 563–4
– spoken corpora 642, 643
– suprasegmental phonology 107, 201
– phonotactics 88, 101
phylogenetics 72
Piaget, Jean 224
picture naming tasks 659
picture-word interference task 662
Pidgin Ngarluma 71
Pidgin Ojibwe 70
pidgincreoles 104
Pidgin-English (Cameroon) 277
pidgins
– anthropology 684
– artificial languages as 131
– colonial studies 697, 701
– contact-induced change 19
– contact-induced syntactic change 26
– continuum of pidgins-creoles 100, 334
– domains 538
– elaborated/expanded pidgins 104, 105
– formation of 101–2
– and globalization 335
– linguistic boundaries 278
– linguistic properties 106–8
– morphology 69–71
– origin of term 100
– pidginization 100–12, 120, 543
– urban-rural contact 325
Pietikäinen, Sari 422, 574, 577, 578, 582, 583, 670, 756, 758, 760
Pike, Kenneth 589
Piniq 594
Pinto, Maria A. 224
Pinyin 345
Piron, Claude 129
pitch accent 476
Pivneva, Irina 216, 217
pivot-matching 150
pixelated society 7
place, discourses in 577–8 *see also* geolinguistics; landscapes, linguistic
place identity 286, 311–12
Plag, Ingo 70, 103
plain language 389–90
planned languages 125
planning, language 56, 153, 345, 357–70, 542, 699
plantation societies 26, 104, 105, 310, 696, 701
Ploog, Katja 271
plural markers 65, 67, 68, 69, 70, 150, 155
pluralism 362, 375, 381, 583, 688, 739
pluricentric languages 392
Plurilingua 3
plurilingualism 140, 229, 257, 715 *see also* multilingualism Pohnpeian 700
Points, Kathleen 314–15
Polian, Gilles 43
policy, language 357–70, 370–84 *see also* law and language
– bottom-up language policies 361, 364, 390
– critical discourse analysis 492
– discourse analysis 630–2, 634
– domain analysis 537, 540, 541–2, 545–6
– education 406, 721
– ethnography 589, 590
– ideologies 299, 688
– language economics 709, 710, 716
– language maintenance 457
– linguistic boundaries 276
– linguistic landscapes 433, 434
– and the media 421, 424, 760
– multilingual contexts 263
– nexus analysis 577
– standardization 389
– translation 789–90, 799
– unofficial 310, 590
Polish 115, 349
politeness 136, 138, 327
political boundaries 274–6 *see also* nation states
political science 765–77
politics, language 357–70
polls 528
polydisciplinarity 791
polyglossia 263, 266, 279
polyglottism 344, 346
polylanguaging 6, 160, 165, 337
polylingualism 140, 267, 324 *see also* multilingualism; plurilingualism
polyphony 240
polysemy 39, 40, 41, 42, 43, 46, 101
polysemy calquing/copying 43
polysystem theory 793
Poplack, Shana 19, 116, 138, 148, 153, 154, 159, 160, 161, 162, 164, 250, 471, 482, 483, 604, 643
Portugal 394

Portuguese 83, 93, 94, 114, 201, 542
positioning theory 629
possessive constructions 68
post-colonial contexts
– anthropology 618, 682
– code-switching 163
– colingualism 323–4
– corpus linguistics 644
– English 333
– gender 248
– language attitudes 239
– and linguistic nationalism 350–1
– and multilingualism 266
– one nation-one language 301
– orthographic reforms 83
– postcolonial studies 695–6, 702
– social network analysis (SNA) 568
– translation 794
post-creole continuum 334
postmodernism 313, 542
postpositions 18, 19
poststructuralism 582, 583, 684, 685, 688, 728
post-threshold second language learning 176, 178, 179
Potter, Jonathan 239, 240, 628
Potts, Deborah 321
power dynamics
– alternative legitimacies 684
– attitudes 235
– biopower 721
– capillary power 630
– code-switching 630
– and colonialism 350, 695, 696, 697, 699, 702, 727
– critical stances 627
– discourse analysis 626–7, 630
– education 413, 721
– ethnography 588, 595
– Foucauldian discourse analysis 626–7
– gender 248, 249
– and language ideology 235, 298, 300, 302, 309, 688
– Language Law 733
– language maintenance 455
– language policy and planning (LPP) 358, 359, 360
– legal studies 737
– and the media 419
– minority language status 372
– multilingual contexts 263, 766
– political science 765–77
– pragmatics 142
– and race 284
– research ethics 668, 675–6
– and speech communities 310
– translanguaging 726–7
– translation 794
Praat 643
practice, nexus of 575
pragmatics 136–47
– code-switching 164
– contact-induced syntactic change 32
– creoles 103, 108
– cross-linguistic influence 137
– Early Second Language Acquisition (ESLA) 191
– mobility 138, 142–3
– pidgins 102
– and semantic change 43
– semanticization of 39
– third language acquisition 226
– translation studies 798, 799
Pratt, Mary Louise 743, 747, 790
Pray, Lisa 192
preference procedures 654–5
prehistoric contact-induced change 13, 14
Preisler, Bent 402
prejudice 238
pre-migration capital 193, 326
preposition stranding 26
prepositions 18, 19, 26
preschoolers 173–4, 189, 190
prescriptivism 56, 127, 132, 299, 301, 302, 540, 790
present participles 68
preservation, language 350 *see also* maintenance of languages
prestige
– code-switching 163
– covert prestige 114, 327
– ethnolects 287
– koines 278
– and language attrition 203
– and language ideology 301–2
– mixed languages 325
– multilingual contexts 262–3
– as reason for borrowing 54, 58, 152–3
– and speech communities 310
– superstrate influence 334
– urban-rural contact 327
– varieties in contact 113, 114
– written language 79
prestige planning 360, 363, 364

Preston, Dennis R. 289
presupposition 154, 155
pre-verbal markers 106
Primary Language Data (PLD) 29
priming 211, 643, 658, 661–2
Prince Edward Island French 26
principal components analysis 481
Principle of Debt Incurred 671, 672
Principle of Error Correction 670–1, 672
Principle of Linguistic Gratuity 671, 672
Prinsloo, Mastin 333, 335
printing 58
Prior, Anat 216, 217
privacy concerns 426, 520–1, 642
probability samples 501, 515, 517, 518, 519
process of language change (versus outcome) 172
processability constraints 29, 32, 33
process-centered approaches 180, 187, 198
processing speeds 213, 214–15, 216
Proctor, Lavanta M. 689
pro-drop 645–6
proficiency
– and code choice 545
– cross-linguistic influence 211
– definitions of 213
– education 409, 411
– measures of multilingual 227–8, 414
– metalinguistic awareness 225
– and research interviews 605
– second language acquisition 213–15, 216
– self-reported 503
– translanguaging 414
– uninstructed language acquisition 186, 190, 193, 194, 199
progressive demasking techniques 658
pronouns 32–4, 71, 149, 251
pronunciation 175, 249, 778 *see also* accent; phonology
proper names 42
prosody 88, 390, 642
Protocol to Ensure Language Rights 734
prototypes 39, 56, 60
Provençal 55
psycholinguistics
– adolescents 176–7
– adult speakers 177–8
– children 174–5
– code-switching 160, 166
– corpus linguistics 643
– experimental psycholinguistics 468
– individual variation in bilingual lexical processing 210–22
– and language attrition 203
– language awareness 412
– lexical borrowing 60
– metalinguistic awareness 228
– multi-competence frameworks 140
– older people 179
– research methods 653–67
– Sapir-Whorf hypothesis 685
– semantic change 44–6
– social network analysis (SNA) 564, 565
– surveys 505
psychotypology 227
puberty 174–5, 176
publics 297, 301–2
Pujolar, Joan 583
Punjabi/Panjabi 352, 493, 686
purism, linguistic 280, 299, 301, 349, 362, 364, 389, 420, 421–4, 457, 698, 760
Purnell, Thomas 91
Purschke, Christoph 240
Putnam, Michael T. 89
Py, Bernard 140, 177, 202

Q

qualitative research 487–500
– anthropology 689
– attrition 203
– correspondence analysis 475–8
– data elicitation and analysis 487–500
– of ethnicity 288
– fieldwork methods 489–92
– gender 249, 250
– language attitudes 236
– linguistic landscapes 437
– nexus analysis 583
– qualitative interviewing 603–4
quantitative research 468–87
– corpus linguistics 643–4
– domain analysis 545, 546
– ethnography 591
– geolinguistics 529
– interviews (research method) 603, 604
– linguistic landscapes 437
– social network analysis (SNA) 562
Quatember, Andreas 516, 517, 518, 520, 521, 522
Québec 301, 363, 377, 378, 380, 408, 709, 713, 770, 779, 784

Quechua 19, 64, 278, 479, 549, 555, 592, 609, 644
queer theory 542
question formation 28, 32–4, 138, 251
questionnaires 138, 202–3, 236, 425, 444, 446, 488, 529–30, 545
Quichua 592–3
Quijano, Aníbal 727
quota sampling 518

R

R (R Development Core Team) 484
Rabus, Achim 79–80
race 284, 615, 616, 617, 778
Radebe, Zodwa 619
radical creoles 104, 105
radio 422, 423, 758
Raimy, Eric 91
Rallis, Sharon F. 590, 591
Ramallo, Fernando 422
Ramani, Esther 593
Rampton, Ben 142, 143, 164, 177, 178, 239, 240, 279, 285–6, 322, 323, 324, 333, 334, 338, 489, 490, 491, 492, 493, 494, 613, 686, 687, 690
random forests 468, 483–4
random sampling 502, 515, 518
ranges of reference 42
rap songs 426
rapid serial visual presentation (RSVP) 659
raplish 336
Rappaport, Joanne 620
Rapport, Nigel 311
Rath, Brigitte 745
Rational Choice Theory 540
Rbrul 484
reaction times (RTs) 657, 658, 660
re-adaptation 201
reading 213 *see also* literacy
real-time processing 200
real-world knowledge 39, 198, 226
Recipient Language (RL) agentivity 25–6
reclamation, linguistic 590 *see also* revitalization of languages
recognition politics 769–70
recording, data 490–1
reductions 31, 32, 101, 102, 120, 278
redundancy 101
reduplication 106–7
Reeves, Lauretta 60
reflexivity 223, 271, 279, 303, 595, 596, 614, 616, 618, 634
refugees 249
refunctionalization 117
refusal 322–3
regimes, language 300–1
regional dialect change 289
regional dialect levelling 113, 117–18
regional varieties 103, 112–23, 275, 286, 289, 321, 345, 374, 391, 528, 733, 758 *see also* dialect variations; indigenous languages; minority language groups
register 113, 137, 419, 420, 421, 422, 483, 538
relativity, linguistic 685
re-learning 204
relegitimization 710
relexification 54, 102, 162, 278
Relexification Hypothesis 162
religion 79, 80, 82, 248, 285, 344, 446, 696, 699, 700, 767
repertoires, linguistic
– in censuses 527
– code-switching 164–5
– complex diversity 772
– domains 542
– education 413, 725–6
– emergence as trend 6
– fluid multilingualism 339
– identity 287
– ideologies 301
– individual 140
– language policy and planning (LPP) 364
– language rights 371
– linguistic capital 633
– and migration 759
– nexus analysis 583
– political science 774
– and research ethics 677
– spatial repertoires 401
– speech communities 173, 310
– superdiversity 336, 687
– translanguaging 410, 413
– and translation 794
– uninstructed language acquisition 187
replacive borrowing 54
replication 44, 53, 149, 150
representativeness of samples 514–23
Research Center on Multilingualism 3
researcher presence 507–8, 579, 580, 602, 615–16, 618–19, 676–7
Resígaro/Risígaro 644

resistance and language attitudes 164, 303, 324, 351, 373, 374, 411, 412
resistance to being researched 618
resistance to language reform 80, 82
resistance to translanguaging 728
resonance 211
response speed measurements 657–60
restandardization 349
resultative constructions 114, 115
retention 101, 103, 175, 203
retrieval, language 505–6, 508, 661
reversal 204
Revised Hierarchical Model 211
revitalization of languages
– anthropology 682–3
– borrowing 153
– campaigns, language 263
– constructed languages 125
– education 725
– ethnography 591
– ideologies 299
– and language maintenance 457
– language politics 362, 363
– linguistic landscapes 436
– media 426
– nation states 346, 348, 350
– politics 769
– social network analysis (SNA) 568
– standardization 393
– translation studies 793–4
Reyes, Angela 690
rhoticity 201
Rice, Keren 675
Rickford, John R 162, 194, 283, 284, 285, 310
Riehl, Claudia M. 204 Rietveld, Toni 482
rights, linguistic 362–3, 370–84, 723, 733, 734, 735–6, 760, 771, 773
Rights and Obligations (RO) sets 163
Rindler Schjerve, Rosita 2, 352, 399
Risígaro/Resígaro 65
Robbins, Lionel 708
Roberts, Ian 27, 29
Roberts, Sarah 70
Rokkan, Stein 767–8
ROLE 539–40, 620
role assignment strategies 655
Romaine, Suzanne 164, 258, 262, 266, 326, 723, 732, 735, 738
Roman Empire 261, 344
Roman script 79, 80, 81, 83, 432–3
Romani 68, 148, 149, 150, 152–3, 154, 552, 556, 644, 646
Romanian 18, 348, 759
Romansch 712
Romero-Little, Mary Eunice 590
Ronjat, Jules 187
Rosa, Jonathan 406
Rosenbaum, Yehudit 432
Ross, Malcolm 44, 45, 47, 69, 697, 698
Rosselli, Mónica 215
Rossman, Gretchen B. 590, 591
routinization 43, 56
Rubin, Donald B. 521
Rubin, Joan 358, 538, 721
Ruíz, Richard 362, 411
Ruiz Vieytez, Eduardo 731, 732, 737
Rumelian Turkish 150
Rumsey, Alan 43, 297, 668
runic scripts 79
rural areas 319–32, 351, 455
rurban areas 321
Russenorsk 101, 102, 105
Russian
– Baltic states 362–3
– borrowing 59, 150
– domain analysis 540, 549
– and Esperanto 129
– and globalization 352
– media 760
– morphology 66, 69
– orthography 77, 84
– in Ukraine 366
Russian republics, former 81, 549, 555
Ryan, Ellen B. 236
Ryan, Louise 565
Ryukyuan 457

S

Sachdev, Itesh 164
Safont-Jordà, Maria-Pilar 137, 226
Safran, William 349
Sakel, Jeanette 13, 40, 64, 66, 114, 155
Sakha 65
Sall, Thierno Seydou 618, 620
Sallabank, Julia 307, 309, 310, 550, 568, 674
Salmons, Joe 91, 93, 94
Salterelli, Maria 204
Sambre, Paul 139
Samek-Lodovici, Vieri 32

Sámi 373, 374, 393, 424, 552, 577, 754–5, 756, 757, 758, 760
sample analysis 495–7
sample representativeness 514–23, 528, 529
sampling errors 502
sampling frames 502, 520
San Andres Creole 424
Sankoff, David 148, 159, 162, 471, 482, 483
Sankoff, Gillian 47, 102, 112, 173, 180, 204, 442, 447, 684
Santa Barbara Corpus of Spoken American English 640
Sapir, Edward 615
Sapir-Whorf hypothesis 685
Saraguros 593
Sardinian 552
Saunders, William 192
Saussure, Ferdinand de 745, 746, 748
Saville-Troike, Muriel 399
Saxena, Mukul 540
Sayahi, Lotfi 248
scalar approaches 577
scale theory 142
scalograms 478–9
Scandinavia 348, 399, 767 *see also specific countries*
Scandinavian 44, 57, 392
scarce resources, and language economics 708
scatterplots 474, 476–8
Schell, Martin 323, 324
Schieffelin, Bambi B. 297, 309
Schiessel, Jamie L. 413, 414
Schiffman, Harold 309, 310, 358
Schiffrin, Deborah 137, 626
Schilling, Natalie 669, 678
Schleef, Erik 251, 252
Schleyer, Johann Martin 128
Schmid, Carol L. 357
Schmid, Monika 179, 180, 199, 200, 202, 203, 204
Schmidt-Rohr, Georg 1, 398, 537, 541
Schmitt, Elena 174–5
Schneider, Edgar 336, 386
school languages 191–2, 227, 719–30
Schooling, Stephen J. 568
science, language of 541
Scollon, Ron 434, 435, 573, 574, 575, 576–7, 578, 579, 582, 583, 584
Scollon, Suzie 573, 574, 575, 576–7, 578, 579, 582, 583
Scotland 445, 454, 474, 761
Scots/Scottish Gaelic 249, 364, 748–9, 760–1
Scott, John 540, 563
script choice 77, 78, 210, 216, 300, 345, 388, 433, 594
script reforms 80–2
Sebba, Mark 65, 68, 77, 78, 80, 82, 84–5, 300
Sebjan- Küöl Even 65
secessionism, linguistic 538
second dialect acquisition (SDA) 113, 117–18, 120
second language acquisition *see also* multilingualism; third language acquisition
– additive bilingualism 784
– adolescents 176
– adult speakers 192–3
– authenticity 302
– 'bottleneck' hypothesis 31, 32, 33
– children 174, 189, 192
– code-switching 160, 165
– contact-induced syntactic change 27–34
– creole genesis similar to 102
– education 720, 723–4
– group second language acquisition 177
– imperfect learning leading to contact-induced change 18–19
– Interface Hypothesis 29
– language attitudes 238
– language economics 710, 712
– learner corpora 641
– linguistic landscapes 434
– metalinguistic awareness 224–7
– monolingual second language programs 725
– morphology/semantics interface 32
– phonology 89
– pidgin genesis similar to 101
– post-threshold second language learning 176, 178, 179
– psycholinguistics 656
– regional features 289
– semantics 45
– syntax/pragmatics interface 33
second language decay 179
Second Life 426
secret languages 126
Seifart, Frank 15, 16, 17, 65, 114, 156, 646
selection malfunction hypothesis 19–20
self-paced reading tasks 659
Selinker, Larry 138, 165
semantics

– artificial languages 130
– borrowing 16, 53, 149, 152–3
– contact-induced semantic change 38–52
– contact-induced syntactic change 31
– contamination 44
– convergence 45
– creoles 103, 108
– discursive practices 46
– fields 60
– individual variation in bilingual lexical processing 211, 212–13
– language planning 699
– neuroimaging 663
– priming 661–2
– semantic change 6
– semantic differential scales 237
– semantic extension 41, 57, 102, 150
– semantic narrowing 42, 59, 102
– semanticization of pragmatics 39
– shift 58, 59
– transparency 102, 155
– uncertainty and the likelihood of borrowing 20
semasiological change 54, 60
semi-creoles 26
semi-languages 126
semilingualism 722–3
semiotics 136, 140, 297, 298, 574, 582, 583, 594–5, 746
semi-speakers 448
Senegal 311, 618, 684
Seneweb 265
sense model 211
sentence-final particles 251
Separate Development Hypothesis 189
separatism, linguistic 347, 351, 363, 373
Serbo-Croatian 80, 274, 347, 349, 392
serial verb constructions (SVCs) 107
Serratrice, Ludovica 32
Sessarego, Sandro 37–4
Seuren, Pieter 26, 71, 106
sexism 246
Shabibi, Maryam 151
Shafto, Meredith A. 178
Shah, Sheena 458, 459, 460, 461, 462
shared norms 309–10, 313, 683–4
Sharma, Devyani 603, 609
Sharpe, Christina 616
Shaws Road (Bothar Seoighe) Gaeltacht 264
Sheltered Content Instruction 409, 414
Shen, Xingjia Rachel 645
Sheng 178, 322, 324–6
Shenk, Petra Scott 686
Shepard, Carolyn A. 780
Shetland Norn 556
Shi, Xingsong 249
shift, language 441–54
– adolescents 177
– anthropology 683
– attitudes 235
– bilingualism 263
– bilingualism without diglossia 538
– code-switching 162
– colonial studies 697
– definition and causes 442–3
– domains 541
– ethnography 590, 593
– functional shift 541
– gender 248, 249
– and language boundaries 275
– legal regulations 732
– lexical items 58
– and linguistic purism 299
– versus maintenance 454, 689
– and the media 422
– minority language status 372
– multilingual contexts 260
– nexus analysis 578, 583
– non-linearity of 275
– pre-modern empires 344
– quantitative research 478
– reversal 363, 725 *see also* maintenance of languages; revitalization of languages
– social network analysis (SNA) 565, 566, 568
– Sociolinguistic Interview (SLI) 603
– syntactic change 26
– unstable bilingualism 310
– urban-rural contact 326
Shin, Naomi Lapidus 250, 252
shipwork contexts 105
Shohamy, Elana 240, 358, 360, 414, 431, 434, 436, 594
Shona 421
Shook, Anthony 211
Shuster, Evelyne 668–9
Si, Aung 646
Siegel, Jeff 102, 103, 112, 117, 120
Sierens, Sven 414
sign languages 4, 210–11, 215, 217, 262, 267, 312, 640, 643 *see also* bimodality
signs (street) 431, 432, 433, 435, 437, 594, 595
Siiner, Maarja 361

Silva Corvalán, Carmen 19, 267, 448
Silverstein, Michael 45, 251, 296–7, 310, 496
Simard, Daphnée 223, 224
Simon, Ellen 94
simplification 101, 116, 120, 128, 162, 176, 178, 200, 278, 335, 419, 448
simultaneous interpreting 661
Sindhi 351
Singapore 263, 301, 457, 460
Singapore English 263, 484
Singh, Rajendra 64
Singler, John Victor 100
Singlish (Singapore Colloquial English) 263
Sinhala 69
Sinti Romani 149
situational code-switching 164, 489
situational context 240, 489, 492, 493, 508
skills, language, economic value of 709–10, 714
Skutnabb-Kangas, Tove 376, 379, 541, 722, 732, 735, 737, 739
Slabakova, Roumyana 31, 32, 33
slang 324
slavery 26, 29, 105, 543, 696
Slavic languages 58, 473, 483–4
Slovak 115, 119, 348, 349
Slovenian 347, 566
Smakman, Dick 455
Smit, Ute 236, 237
Smith, Benjamin 550
Smith, Jennifer 31
Smith, Linda Tuhiwai 618
Smith, Norval 107
social actions (nexus analysis) 573–87
social boundaries 279
social capital 138
social categorization 235, 237, 240, 251–2, 285, 308, 313–14
social class 236, 239, 250, 252, 564, 565, 617, 686
social constructivist approaches 239, 490, 684
social constructs
– age as 172–84
– communities as 311
– identity as 685
– languages as 140, 309
– and linguistic nationalism 346–8
– race as 284
– and sample analysis 495–7
– of space 434
– speech communities 309
social context
– bilingual child language acquisition 190
– code-switching 159
– and language attrition 199
– pidgins and creoles 104–8
– as reason for borrowing 152–3
– social cues 336
social hierarchies 309, 310, 689, 727
social identity marker, language as 19, 778–85
Social Identity Theory (SIT) 235
social indexicality 163, 234–45, 247, 249, 251, 287
social justice 596
social meaning 78, 240, 446–7, 683
social media
– as data source 491
– media linguistics 419, 425–6, 760–1
– multilingual contexts 265
– online communities 312
– and social networks 562
– sociopragmatics 138
– as source of participants 610
social media method (SMM) 425
social motivations approaches 67
social network analysis (SNA) 561–73
social networks
– code-switching 545
– communities of practice 264–5, 314
– definition of 562–3
– and differences in learning languages 194
– domains 540
– and gender 250
– identity 249
– language economics 711
– and language shift 275, 445–6
– languages as ligatures 775
– large scale studies on 180
– Network Strength Scale 565
– open versus closed 540
– and social indexicality 251
– sociopragmatics 138
social practice theory 313
social psychology 203, 235, 238, 239, 240, 384, 562, 628, 777–88
social status, marking 117
social values 248, 249, 711, 712
socio-cognitive processes 297, 563, 798
sociocultural factors
– across the lifespan 172
– contact-induced syntactic change 26, 34

– creoles 26
– education 406, 720
– lifestyle diglossia 540
– migration 333
– participant observation 613, 614
– pragmatics 136, 137–9
– semantic borrowing 45
– semantic change 48
– social psychology 778
sociodemographic data 444
socioemotional factors 194, 203
sociohistorical practices
– creole genesis 102
– and language attrition 204
– lexical borrowing 59
– participant observation 616
– pidgins and creoles 100, 104
– semantic borrowing 46
sociolinguistic factors
– adolescents 177
– adult speakers 178
– age 173
– bilingual communities and dialect variation 113, 121
– bilingualism 210
– borrowing 19
– children 175
– code-switching 159, 163–5
– domains 397, 399–400
– ethnicity 283, 286–8
– gender 247, 249, 252
– integration of loans 59
– language attitudes 236, 239
– language awareness 412
– language ideology 297, 337
– linguistic landscapes 431
– loanword acceptance 55
– and the media 420, 427
– metalinguistic awareness 228
– minority groups 372–5
– naturalism 302–3
– older people 179–80
– and pragmatics 136
– prestige and borrowing 54
– script choice 79–80
– speech communities 308
– superdiversity 267
– supervernacular 338
– third-wave sociolinguistics 240
– and written language 77
Sociolinguistic Interview (SLI) 601–10
sociological research approaches 488–9
sociometry 562
sociopolitical agency 7
socio-political factors *see also* policy, language
– creoles 108
– dialects 113
– language economics 709
– and language ideology 298–9, 302
– language politics, policy and planning 357–70
– linguistic purism 299
– and the media 419
– nexus analysis 576
– pidgins 101
– political science 765–77
– states, nations and language contact 343–56
– written language in periods of unrest 80
sociopragmatics 137, 138, 166
socio-spatial contexts 524
software, data analysis 469, 482, 484
solidarity-marking codes 163, 164, 324
Solomon Islands Pijin 104
Sorace, Antonella 32, 33
Sorani Kurdish 149
Sorbian 473, 645, 646
Sorensen, Arthur 442
Sotho 277
Soukup, Barbara 437
sound system 91–2 *see also* phonology
Source Language (SL) agentivity 25–6
South Africa 277, 320, 322, 323, 337, 371, 458–9, 462–3, 543, 593, 605, 736
South Sea Pidgin English 698
South Slavic 347
South Tyrol 224, 228–9, 230, 526, 527, 528, 760
Spanish
– *Abstand* and *Ausbau* 347
– adolescents 176
– borrowing 19, 27–34, 114, 153, 154
– in Canada 603, 609
– and Chicano English 289
– as colonial language 261, 344, 699, 700, 702
– corpus linguistics 644
– gender stratification 250
– and globalization 352
– ideology of anonymity 303
– languages of wider communication (LWCs) 442
– mixed languages 278

– morphology 64, 66
– multilingual contexts 260
– New Mexican Spanish 482, 645
– official status languages 771
– pragmatics 137–8
– syntactic borrowing 27–34
– in the US 263, 275, 423, 446, 552
– sparse data 469
spatial contexts 140, 142, 431–41, 524, 576, 577–8
spatial repertoires 401
spatio-temporal scales 287–8
speaker numbers/populations 262, 393, 443, 445, 454, 456, 739
speech accommodation theory 235
speech acts 137
speech communities 307–19
– attrition 198, 204
– code-switching 162, 165
– defined 308
– hybrid 10
– linguistic landscapes 434
– media/communication studies 760
– multilingual contexts 267
– nexus analysis 575
– older people 180
– semantic change 43
– shift 445
– shift away from concepts of 7, 165, 312–14
– varieties in contact 113, 120
speech recognition technology 641, 797
Speelman, Dirk 60
spelling 77, 80, 388, 426, 495, 496 *see also* orthography
sphericules 755
Spitulnik, Debra 684
Spitzer, Leo 39
Spitzmüller, Jürgen 626, 632, 633, 634
Spivey, Michael J. 657
Spocanian 129
spoken corpora 640, 642
spoken language translation 796
Spolsky, Bernard 309, 310, 358, 361, 362, 433, 541, 542, 589, 698, 702, 737, 738
spontaneous speech, eliciting 604–5, 607
sprachbund 115
Sprachvolk 308
SPSS (Software Package for the Social Sciences) 484
Sranan 27, 30
Sranantongo 106
Sri Lanka 411, 593
Sri Lanka Malay 69
Stær, Andreas 495
Stähli, Adrian 550
stance-taking 7, 246, 247, 249, 251, 252, 606
standard language, ideology of 297, 299, 302, 390, 537, 609, 688, 722, 755, 756, 768
standardization 384–97
– constructed languages 128
– domain analysis 542, 543
– educational sciences 722
– ideologies 298
– and the media 419, 421–2, 426
– nation states 343, 346, 349
– and the nation-state 445
– nexus analysis 582
– politics 359, 362, 364, 767, 769
– urban-rural contact 321
Starý, Zdeněk 2
statistical analysis methods 444, 468–87, 515–17, 521, 545, 546, 565
status planning 359, 360, 364, 721
Stebbins, Tonya N. 615
Steffensen, Sune Vork 746
Steger, Manfred B. 309
stereotypes 235, 237, 246, 250, 300, 493, 506, 782, 783
Stevens, Kenneth Noble 92
Stewart, William 553
stigmatization 46, 263, 301, 410, 445, 530, 588, 688, 719
Stockberger, Inge 584
Stockhammer, Robert 745
Stocking, George W. 617
Stoessel, Saskia 568
Stolz, Christel 16, 155
Stolz, Thomas 16, 155
Storch, Anne 620, 698
Storey, John 617
strategies, learning 226, 230
Strauss, Anselm L. 494
street language styles 324, 325, 327, 495–6, 564
Strevens, Peter 259
Stroop tasks 216, 217
Stroud, Christopher 333, 335, 454, 456, 555
structural borrowing 150–2
structural boundaries 278
structural constraints theories 64, 67
structural integration 154

structural interference 701
structural transfer 25
structuralism 126, 252, 539, 722, 726, 744
Stuart-Smith, Jane 420, 423, 427
style-shifting 162, 165, 310, 314
stylistic variation 57, 58, 113, 117, 160, 164–5, 240, 325, 448, 606
Sub-Commission on Prevention of Discrimination and Protection of Minorities 378
subjective vitality questionnaire 444
subject-verb agreement 28, 31–2, 68, 71, 106
subject-verb inversion (questions) 33–4
substitutions, morphological 175
substitutions, phonological 88, 93
substrate influence 13, 27, 45, 58, 101, 102, 103, 106, 334, 447
subtitling 422, 797
subtractive bilingualism 407, 725, 783–4
suffixes 56
Sumerian 344
superdiversity 332–43
– anthropology 687
– code-switching 165
– emergence as trend 6, 7
– ethnography 592
– identity 287
– language rights and policies 375, 381
– literary studies 749
– media/communication studies 755
– migration 333
– multilingual contexts 266–7
– nexus analysis 583
– politics 771–2
– pragmatics 142
– speech communities 310
superstrate influence 101, 102, 103, 334
supervernacular 338
supralexification 54
suprasegmental phonology 107, 201
Suriname Creoles 104, 106
surveillance 617
surveys 500–13, 525, 530 *see also* interviews (research method)
Sutherland Gaelic 249
Svendsen, Bente Ailin 175, 177, 277, 568
Swadesh, Morris 57
Swahili 153, 154, 324, 325, 360, 460, 549, 644, 699
Swain, Merrill 165, 408, 410, 784
Swales, John M. 312
Sweden 262, 541, 594, 759
Swedish 64, 262, 287, 348, 380, 392, 398, 527, 541, 594, 759, 760, 768
Sweetser, Eve 39
Swigart, Leigh 164, 684, 690
Switchboard Corpus 640
Switzerland 275–6, 349, 363, 398, 431, 501, 526, 537, 550, 552, 628–9, 711–13
symbolic categories 91
symbolic power 299
symbolic values
– attitudes 235
– boundaries 271, 275
– gender 248
– language as social identity 778–80
– language economics 708, 713–15
– language policies 360
– and language shift 445
– legal regulations 732
– linguistic landscapes 432
– script choice 80, 82, 85
synecdoche 42
synonyms 58, 389
syntactic change 24–38, 40, 41, 43–4, 46–7, 55
syntax. Minimalist 6, 161, 643
syntax/discourse interface 27
syntax/pragmatics interface 32–3
syntax/semantics interface 30
synthetic languages 125
Syriac 344
System Morpheme Principle 68
systemic imposition 93–4
Szmrecsanyi, Benedict 484

T

taboo forms 324
Tabouret-Keller, Andrée 162, 279, 310, 314, 386, 441, 445, 684
tacit knowledge 229
Tadmor, Uri 16, 42, 57, 152, 699
Taeschner, Troute 165
tag questions 251
Tagalog 91, 93, 95–6, 186
tagging corpora 642
Tagliamonte, Sali 483, 601, 605
Taiwan 594
Tajfel, Henri 235, 779, 781, 782, 784
Takia 47
TAM (tense, aspect, mood) inflection 65, 69, 71, 106

Tamil 69, 263, 351, 411, 459, 460, 593
Tanzania 360, 461–2
target populations 502
Tariana 41, 448
task-based learning 411
Taylor, Charles 769
Taylor, Talbot J. 297
Te Reo Māori 424
teacher education 412–13, 593
teaching 175, 723–4
Tel Aviv 436
television 419–20, 423, 424, 458, 756, 758
tense 43, 46, 70, 106, 130
Teochew 263
Terluin, Ida 321
terminology of the field 6–7
territoriality principle 735
text frequencies 20
Thai 434, 435
Thailand 434
thematic roles 67, 655
Theory of Norms and Exploitations 60–1
thick descriptions 613
Third Grammar 161
third language acquisition 226–7
third spaces 413, 687
third-wave sociolinguistics 240, 613
Thomas, George 115
Thomason, Sarah Grey 13, 14, 15, 18, 24–5, 47, 54, 57, 64, 65, 66, 100, 112, 113, 114, 115, 148, 152, 155, 156, 159, 162, 172, 175, 278, 420, 542, 676, 684, 697, 701, 703
TH-stopping 286, 287
Thurlow, Crispin 434
Ticca, Anna Claudia 141
Tifinaġ 82
Tigrinya 83, 435
time/temporality 576, 577–8, 583
Titone, Renzo 224
Tiwa 479, 480
Tok Pisin 102, 104, 260, 445, 615
tokenism, language 436
Tokyo 432, 434, 582
Tollefson, James W. 358, 361, 362, 445, 590
TOPIC 539–40
topic-continuity 32
Toronto 288, 289, 608–9
Torres Cacoullos, Rena 482, 609, 645
Torres Strait Islanders 373
Tourangeau, Roger 503
tourism, linguistic 747
Tovares, Alla 325, 327
Town Bemba 684
trading contexts 105, 709–10
transcriptions, data 494, 592, 641–3
transcultural flows 311, 324
transfer *see also* borrowing
– borrowing versus imposition 25–6
– code-switching 159, 161–2, 165
– to colonizer language 700
– education 720, 724
– and language attrition 199, 204
– and the media 419
– metalinguistic transfer 229
– paradox of transfer 227
– phonology 88–9, 93–4
– predisposition of forms to 58
– semantic change 45
– syntactic change 24–5
– translanguaging 726
translanguaging
– children 175
– and code-switching 160
– education 407, 410, 413–14, 726–7
– emergence as trend 6
– fluid multilingualism 337
– language policy and planning (LPP) 364
– linguistic landscapes 436
– linguistic repertoires, individual 339
– multidisciplinarity 9
– pragmatics 140
– and speech communities 310
– and superdiversity 165, 267
– urban-rural contact 322
translation 39, 141, 265, 641, 642, 660, 697, 746, 750–1
translation pairs 211
translation studies 421, 425, 641, 750, 789–801
translingualism 337, 747, 748, 795
transliteration 84
transnational communication 311, 419, 583, 687, 690, 771, 773
transnationalism 688–9
trans-super-poly-metro movement 337
Trask, Robert Lawrence 57
Traugott, Elizabeth C. 39, 40, 743
trauma 737
Traverso, Véronique 141
Travis, Catherine E. 609, 645
Trechter, Sara 669, 672, 673, 675, 677
tree diagrams 480–1

Treffers-Daller, Jeanine 115, 120–1, 159, 164, 173, 186, 267
trends, current 5–10
triangulation of data 446, 488, 592, 798
tribal languages 381
Tribur, Zoe 564
triggering hypothesis 166
triglossia 325
trilingualism 191, 214, 228, 265, 301, 352, 758–9 *see also* multilingualism; third language acquisition
Trio 102
Trio-Ndyuka Pidgin 105
Trobriand Islands 615, 618
Troyer, Robert 437
Trubetzkoy, Nikoai S. 128
Trudgill, Peter 18, 101, 113, 116, 117, 118, 162, 172, 176, 178, 246, 275, 321, 386, 392, 399, 524
truth 626
truth claims 155
Tseng, Ming-Yu 139
Tsisipis, Loukas D. 401, 448
Tsotsitaal 322, 324–6, 543
Tsuda, Sanae 422
Tsunoda, Tasaku 556
Tuareg 82
Tucano 448
Tukanoan 41, 44, 46
Tulp, Stella M. 433
Tuominen, Anne 326
Turi, Joseph 731, 732, 733, 735
Turkey 349
Turkic 344
Turkish 18, 53, 55, 58, 65, 68, 81, 82, 150, 154, 352, 543, 644
Turku 70
Turner, John C. 235
Tuxtla Chico 448
Tuyaka 46
Twitter 265, 312, 419, 425, 760
two-way/dual language programs 407
Tyneside 446
typographic mimicry 84–5
typology 549–61
– bare nouns 30
– and borrowing 54–5, 64, 150, 155
– code-switching 161
– constructed languages 128
– contact-induced change 15–16, 20
– contact-induced semantic change 40–3
– contemporary model of 553–7
– creoles 26, 27, 69, 71, 103, 107, 701
– definition of 550
– digraphia 79
– language shift 444
– minority language groups 373
– and orthography 81
– pidgins and creoles 120
– psychotypology 227
– third language acquisition 227
– varieties in contact 112, 113, 114, 310
Tyson, A. F. 424

U

Uchikoshi, Yuuko 189, 194
Uisai 542
Ukraine 366
Umbel, Vivian 191, 194
unbound morphemes 15–16, 18, 65
unconscious phenomena in language awareness 223
under-specification 101
unification, linguistic 362, 698
uniformization 715, 716, 768
uninstructed language acquisition 185–97
United Nations (UN)
– Charter 363
– Convention on the Rights of the Child 733, 734
– Declaration of Human Rights 371, 377, 631
– Declaration on the Rights of Indigenous Peoples 371, 373, 379, 380, 734
– Declaration on the Rights of Persons Belonging to National or Ethnic, Religious and Linguistic Minorities 363, 379, 380, 631, 734
– Human Rights Committee 378
– language policy and planning (LPP) 361
– Minorities under international law 378
– UNESCO scale of endangerment 444, 456, 550, 755
– UNESCO *Use of Vernacular Languages in Education* 721–2
United States (US)
– borrowing 42, 90
– census data 527
– Chicano English 289
– education 406, 413
– English 262, 362, 688
– ethnicity 617

– ethnographic research 589, 590
– ethnolects 286–7
– heritage languages 423, 456
– linguistic nationalism 350
– multilingualism 193, 262, 275, 285, 286, 288, 289, 443, 580–1
– Spanish in 539, 686
Universal Declaration of Linguistic Rights 734
universal experiences 57
Universal Grammar 160, 161
universal grammar 19
universal processes 24–38, 70, 159, 160–1
universality 14, 15
universals 103
unmarkedness 177, 291
unstable bilingualism 26, 310, 461, 538, 770
Urban, Matthias 39, 40, 43, 631
urban areas 263–4, 267, 277, 319–32, 401, 426, 432, 455, 542, 543, 564, 592, 615–16, 771–2, 794
urban minorities 375
urban speech styles 175, 177, 178, 246, 287
urban vernaculars 336–40
Urban Wolof 684, 690
urbanization 566
Urdu 66, 78, 351, 352, 684
Uribe-Jongbloed, Enrique 422, 424, 425, 754
usage planning 359–60, 364
usage-based approaches 178, 279, 363, 756
utility-maximization 326

V

vagueness 138
Vaid, Jyotsna 655
Vail, Peter 556–7
Van Avermaet, Piet 414
van Coetsem, Frans 13, 19, 25, 88, 89, 90, 91, 92, 93, 94, 96, 115, 162, 697
van Doorslaer, Luc 790, 791
Van Hell, Janet G. 210, 211, 212, 214, 215
Van Heuven, Walter J. B. 211, 212, 658, 664
van Hout, Roeland 20, 482
van Leeuwen, Theo 139
Van Mensel, Luk 433–4, 435, 436
Van Parijs, Philippe 732, 738, 769
Vandenbroucke, Mieke 433
Vandenbussche, Wim 364, 385, 386, 426
Vanhove, Martine 40
Vanuatu languages 45, 47, 250
variable rule analysis 482–3
Variationist Linguistic Landscape Study (VaLLS) 437
variationist sociolinguistics
– across the lifespan 173
– code-switching 162
– corpus linguistics 643
– data analysis methods 492
– interviews (research method) 602, 604, 606, 607
– literary studies 744
– and the media 420, 424
– participant observation 613
– quantitative research 468, 482
– social network analysis (SNA) 563, 565
– speech communities 308, 313
– varieties in contact 113
– varieties in contact 112–23, 235, 309, 310–11, 336, 358, 363, 364
Varis, Piia 685, 687
Vasco Uribe, Luis Guillermo 620, 621
Vaupés languages 44, 45, 46, 47
Vehicular Malay 69
Velázquez, Isabel 569
Velupillai, Viveka 26
Venetian 56, 60
Veneto 250
Venuti, Lawrence 750, 789
verbal guise studies 237, 238
verbal hygiene 421
verbal morphology 30
verbalization data 225
verbs, borrowing 154
vernacular speech 604
vernacularization 362, 424, 426
Verschueren, Jef 136, 300
Vertovec, Steven 142, 165, 266, 333, 375, 583, 687, 772
video games 420, 756
video recordings 491, 494, 592, 617
Vienna Convention on the Law of Treaties 378
Vietnamese 81
Vikør, Lars S. 362
Vincze, Laszlo 759
virtual communities 312, 593 *see also* social media
visual perception experiments 434
visual world paradigm 657
visualization of data 471
visuals 435 *see also* graphics
vitality, measures of language 444–5, 550

vocabulary
– asymmetrical borrowing of 16, 19
– and language attrition 200
– mixed languages 324
– pidgins and creoles 107–8
– word formation and vocabulary expansion 52–63
Voegelin, Charles F. 1
Vogt, Hans 159
Voice-Onset Time 92, 201, 469, 472, 482
voicing distinctions 89, 91–2, 93, 94–6
Volapük 128
Volterra, Virginia 165
von Bortkiewicz, Ladislaus 515
vulgarisms 324
Vygotsky, Lew 224

W

Wagner, Daniel A. 77
Wagner, Suzanne Evans 172
Waksman, Soshi 436
Wales 363, 736
Walker, James A. 290
Wallace, Stephen 156
Walser German 550
Walters, Keith 248, 249
Wanderwörter 14
Wang, Xuan 592, 593
Wardhaugh, Ralph 307
Warlpiri 68–9
Warnier, Jean-Pierre 276
Warnke, Ingo H. 626, 632, 633, 634, 695
Warren-Rothlin, Andy 79, 80, 81, 84
Warruwi 260
Washington DC 435
Waskia 47
weaker links hypothesis 213
web surveys 501, 507, 518
Weber, Jean-Jacques 337
Weber, Max 766
Weber-Fox, Christine M. 663
WebMAUS 643
Webster, Noah 350
Wee, Lionel 335, 339
Weinreich, Uriel 1, 2, 5, 15, 16, 18, 40, 44, 57, 58, 64, 65, 88, 138, 139, 148, 149, 159, 173, 198–9, 258, 431, 442, 588, 719, 720, 721, 723, 724, 726, 790, 793
Weisgerber, Leo 308
Welsh 363, 374, 421, 424, 425, 538, 552, 644, 760, 781
Wenzel, Veronika 433
Werker, Janet 189
Wertsch, James 574
Western bias 79, 81, 100, 252, 267, 285, 303, 388, 674, 722, 727, 750, 755, 795
Westernization 79, 81
Wetherell, Margaret 239, 240, 628
we-/they- codes 164
Wetrogan, Signe 509
Whaley, Lindsay J. 309, 312, 446, 550, 555, 682
wh-fronting 33
Whinnom, Keith 71
White, Lydia 27, 31
White, Paul 551
white privilege 285
Whitecloud, Simone S. 674
Whitney, William Dwight 1, 25, 57, 64
Wichí 260
Wichmann, Søren 14, 20
Wide, Elizabeth 756
Wiemer, Björn 150
Wigboldus, Daniël 782, 783
Wilkins, David P. 13, 15, 16, 39
Willensnation 276
Williams, Ann 113, 116, 172, 178
Williams, Colin H. 524, 532
Wilson, James 118, 120
Wilson, Robert 47, 66, 684
Windisch, Uli 276
Winford, Donald 1, 2, 13, 19, 25, 26, 60, 65, 89, 100, 101, 102, 103, 104, 112, 198, 310, 442, 543, 790
Winsler, Adam 189, 194
Winter, Joanne 248
Winter, Werner 58
Winters, Margaret E. 54
Wirth-Nesher, Hana 748
Wissink, Marieke 565
within-participant designs 214
Wohlgemuth, Jan 16, 42
Wölck, Wolfgang 2
Wolcott, Harry 587
Woldemariam, Hirut 431, 435, 436
Wolff, Stefan 732, 737
Wolfram, Walt 514, 520, 671, 672
Wolfson, Nessa 607
Wolof 265, 324, 684, 690
women's speech 246–56, 321, 327
Wong-Scollon, Suzie 434, 435

Woolard, Kathryn A. 297, 300, 303, 309, 323, 385, 687, 688, 689
Woolford, Ellen 161
word formation and vocabulary expansion 52–63
word lists 605
word order 102, 103, 106, 129, 131, 149, 200
word production studies 212, 215
word recognition studies 214
word translation tasks 660
working memory 216, 661
workplaces 138, 177, 248, 250, 337, 424–5, 541
World Englishes 336, 447, 484
world languages 273 *see also* globalization
World Loanword Database 152
worldview, and language 685
Wortham, Stanton 608
Wright, Wayne 719
writing reforms 80–2
writing systems 76–88
written language *see also* literacy; literary studies
– Chinese 345
– colonial studies 697–8, 699–700
– corpora 639–40
– critical discourse analysis 492
– as data source 468–9
– default/unmarked writing systems 79
– digitally-mediated interaction 141
– domain analysis 545
– education 722
– linguistic landscapes 432
– loanwords 56
– media linguistics 420
– orthography and graphemics 76–88
– script choice 77, 78, 210, 216, 300, 345, 388, 433, 594
– script reforms 80–2
– standardization 388, 389, 390, 393
– translation 796, 797
– translation studies 791
Wuorinen, John 768
Wyman, Leisy T. 594

X

Xhosa 337
Xiang, Huadong 664

Y

Yabacrane 324
Yiddish 129, 150, 447, 456, 538, 556
Yılmaz, Gülsen 200, 201, 1999
Young, Andrea 412
Young, Michael W. 618
youth language 175–7, 277, 287, 322, 324, 325, 426, 564, 594, 727
YouTube 265, 426
Yucatec Maya 685
Yugoslavia 274, 349
Yukon Kaska 591
Yungueño Spanish 27–34

Z

Zabrodskaja, Anastassia 436
Zalizniak, Anna 40, 48
Zamenhof, Lejzer 128–9, 131
Zentella, Ana Celia 445–6
Zentz, Lauren 688
zero morphemes 65
Zeshan, Ulrike 640
Zhang, Jingning 423
Zhu Hua 140
Zhuravleva, Anastasia 425, 426
Zimbabwe 321, 323, 421
Zimmermann, Klaus 701
Zufferey, Sandrine 137
Zulu 277, 324
Zuni Indians 614