Constraints in Discourse 2

EDITED BY

Peter Kühnlein
Anton Benz
Candace L. Sidner



John Benjamins Publishing Company

Constraints in Discourse 2

Pragmatics & Beyond New Series (P&BNS)

Pragmatics & Beyond New Series is a continuation of Pragmatics & Beyond and its Companion Series. The New Series offers a selection of high quality work covering the full richness of Pragmatics as an interdisciplinary field, within language sciences.

Editor

Anita Fetzer
University of Würzburg

Associate Editor

Andreas H. Jucker
University of Zurich

Founding Editors

Jacob L. Mey
University of Southern Denmark

Herman Parret
Belgian National Science Foundation, Universities of Louvain and Antwerp

Jef Verschueren
Belgian National Science Foundation, University of Antwerp

Editorial Board

Robyn Carston
University College London

Thorstein Fretheim
University of Trondheim

John C. Heritage
University of California at Los Angeles

Susan C. Herring
Indiana University

Masako K. Hiraga
St. Paul's (Rikkyo) University

Sachiko Ide
Japan Women's University

Kuniyoshi Kataoka
Aichi University

Miriam A. Locher
Universität Basel

Sophia S.A. Marmaridou
University of Athens

Srikant Sarangi
Cardiff University

Marina Sbisà
University of Trieste

Deborah Schiffrin
Georgetown University

Paul Osamu Takahara
Kobe City University of Foreign Studies

Sandra A. Thompson
University of California at Santa Barbara

Teun A. van Dijk
Universitat Pompeu Fabra, Barcelona

Yunxia Zhu
The University of Queensland

Volume 194

Constraints in Discourse 2
Edited by Peter Kühnlein, Anton Benz and Candace L. Sidner

Constraints in Discourse 2

Edited by

Peter Kühnlein
function2form

Anton Benz
Centre for General Linguistics, Berlin

Candace L. Sidner
Worcester Polytechnic Institute

John Benjamins Publishing Company
Amsterdam / Philadelphia

™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Constraints in discourse 2 / edited by Peter Kühnlein, Anton Benz and Candace L. Sidner.
p. cm. (Pragmatics & Beyond New Series, ISSN 0922-842X ; v. 194)
Includes bibliographical references and index.
1. Discourse analysis. 2. Constraints (Linguistics) I. Kühnlein, Peter. II. Benz, Anton, 1965-
III. Sidner, C. L. IV. Title: Constraints in discourse two.
P302.28.C67 2010
401'.41--dc22 2009047143
ISBN 978 90 272 5438 2 (Hb ; alk. paper)
ISBN 978 90 272 8854 7 (Eb)

© 2010 – John Benjamins B.V.
No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Rhetorical structure: An introduction **1**
Peter Kühnlein

Clause-internal coherence **15**
Jerry R. Hobbs

Optimal interpretation for rhetorical relations **35**
Henk Zeevat

Modelling discourse relations by topics and implicatures:
The elaboration default **61**
Ekatarina Jasinskaja

The role of logical and generic document structure in relational
discourse analysis **81**
Maja Bärenfänger, Harald Lüngen, Mirco Hilbert & Henning Lobin

Obligatory presupposition in discourse **105**
Pascal Amsili & Claire Beyssade

Conventionalized speech act formulae: From corpus findings
to formalization **125**
Ann Copestake & Marina Terkourafi

Constraints on metalinguistic anaphora **141**
Philippe De Brabanter

Appositive Relative Clauses and their prosodic realization in spoken
discourse: A corpus study of phonetic aspects in British English **163**
Cyril Auran & Rudy Loock

Index **179**

Rhetorical structure

An introduction

Peter Kühnlein

0.1 General remarks

Texts, and in general types of discourse, vary along a multitude of dimensions. Discourse can be spoken or written, monological or an exchange between a number of participants, it can be employed to inform, persuade (and serve many more or even mixed functions), it can take place in various settings and be arbitrarily extensive. However, some characteristics are shared by all kinds of texts.

One of those shared properties is that text, and discourse in general, is structured, and they in turn are so in a multitude of ways: classical written text as the present, e.g., typically has logical and graphical structuring into paragraphs, sections, chapters etc. depending on the type of text or the genre, but it is by nature monological. Spoken dialogue, at the other end of the spectrum, *inter alia* is characterized by assignment of and changes in roles participants assume in the exchange, stretches of overlapping speech, repairs and many more phenomena that are not regularly observed in written text (notwithstanding chats on the internet and the like) and which give rise to completely different types of structure. All of these forms of communication fall under the common denominator *discourse*; we will keep using this cover term here to refer to them.

The different kinds of structures in discourse have been object toresearch for a considerable time. One type of structure has been of special interests for researchers working in more formal paradigms and has been hotly discussed ever since: it is what is called the *rhetorical* or *coherence structure*. Rhetorical structure is built by applying rhetorical relations recursively to elementary units of discourse. This kind of structure is to be distinguished from, e.g., *cohesive* structure that comes to existence by means of, e.g., coreference in various forms.

The present collection comprises papers that give a wide variety of perspectives on the constraints governing discourse structure, and primarily rhetorical structure: various ways of thinking of constitutive units of discourse along with a variety of conceptions of rhetorical relations are presented, and the issue of which kind of structure is right for the description of the rhetorical make-up of discourse is tackled from different points of view.

Accordingly, this introductory chapter is intended to provide the necessary background to understand the discussions by sketching as briefly as possible the state of the art; the reader is referred to the individual chapters in the volume where appropriate.

As the previous volume, *Constraints in Discourse* (Benz & Kühnlein, 2008), the present one is the result of selecting and compiling papers that are extended versions of presentations at a workshop in the series "Constraints in Discourse." The second of these workshops was held in Maynooth, Ireland, and organized by Candace Sidner (chair), Anton Benz, John Harpur and Peter Kühnlein. All the authors who contribute to the present volume submitted their re-worked and substantially extended papers to a peer reviewing process, where each author had to review two other authors' papers. In addition, John Benjamins conducted an own reviewing process before agreeing to publish the collection. This two-stage reviewing process is intended to secure high quality of the contributions.

0.2 Elementary units

Just as in any formal description of structures, one basic step in describing rhetorical structure of a discourse is to identify the elementary units. Due to the multitude of dimensions along which discourse can vary and due to differences in theoretical assumptions, there is no consensus on what to count as an elementary unit.

Exemplarily, there is a divide between proposals for different domains: a proposal for spoken discourse can refer to intonational features as an important criterion for segment status, whereas a proposal made specifically for written discourse can't. On the other hand, a proposal set up for written discourse can make reference to punctuation and syntactic units, whereas the first is absent and the the latter are not reliably correct in spoken discourse.

Research in prosodic features of discourse and its segments reaches far back: Butterworth (1975) reports that speech rate changes during discourse segments, being higher at the end of a segment than at the beginning. Chafe (1980) observes that pause lengths are varying at segment boundaries too. Much corpus based and computer linguistic research in this area was conducted by Julia Hirschberg with various collaborators, e.g., Hirschberg and Pierrehumbert (1986), Grosz and Hirschberg (1992), Hirschberg and Pierrehumbert (1992), or Hirschberg and Nakatani (1996).

One of the most detailed empirical inquiries into the relation between discourse segmentation and prosody is given in (Hirschberg & Nakatani, 1996),

and the methodology employed there deserves a little closer description. The authors set up a corpus of directives, where subjects had to give route descriptions of varying complexity through Boston. The first series of descriptions was given spontanously by the subjects, recorded and then transcribed. In a second series, the same subjects read the corrected (i.e., freed of false starts etc.) transcripts of their route descriptions, and again the speech was recorded and transcribed. The data obtained from one of the speakers were prosodically transcribed using the ToBI standard, split into INTERMEDIATE PHRASES, pause lengths were measured and fundamental frequencies (F0) and energy (RMS) calculated.

Then two groups of annotators marked up the texts with segment boundaries: one group was given the transcripts only, the other group was given transcripts plus recorded speech. The theory that served as background to segmenting the texts was that of Grosz and Sidner (1986). Their account is potentially independent from domain, i.e., applicable to both spoken and written discourse; Grosz and Sidner claim that discourse structure actually consists of three distinct, but interacting levels. The most central of these levels is the intentional one: for every coherent discourse, that is the claim, one can identify an overarching discourse purpose the initiating participant seeks to pursue. The segments of discourse according to this theory correspond to sub-purposes, the so-called discourse segment purposes. Elementary units in this theory correspond to single purposes. The two other levels, attention and linguistic realization, concern which objects are in the center of discourse and how the discourse is actually realized using cue-phrases and special markers.

The results obtained by (Hirschberg & Nakatani, 1996) in their study on spoken discourse confirm previous findings and reveal much more detail than, e.g., the work by Butterworth (1975); Chafe (1980) reports: both F0 and energy are higher at the beginnings of discourse segments than at their ends. Speech rate on the other hand increases towards the end of a segment, and pause lengths during a segment are shorter than before a segment beginning and after a segment end. So segment boundaries as judged according to purposes indeed seem to be correlated with measurable changes in the speech signal.

These, and similar, findings seem to indicate good mutual support between the intention-based theory of discourse structure developed by Grosz and Sidner (1986) and the claim that discourse segment boundaries are marked intonationally in spoken discourse.

Considerably more work than on spoken discourse has traditionally been devoted to written text than to spoken discourse phenomena. The pioneering work dates back to the 80s, and the cited work by Grosz and Sidner (1986) is among that. An account of discourse structure that had comparable impact at that time was developed by Mann and Thompson (1987a, 1988b). This account,

known as Rhetorical Structure Theory (RST) was explicitly developed as a means to capture analysts' judgements about writers' intentions while composing texts. Thus, RST is devoted to the analysis of written text, but the analysis is not primarily guided by linguistic surface structure: according to its founders, it is rather "pre-realizational" in that it aims to describe the function of (the interplay of) constituents in abstraction from linguistic realization.

Mann and Thompson (1987a, 1988b) claim that the base case for linguistic expressions conveying intentions are clauses of certain types: main clauses, non-restrictive relative clauses are of the right variety, whereas, e.g., restrictive relative clauses and complement clauses (e.g., in subject or object position in a matrix clause) are not counted as minimal units. It turned out in the development of RST since its inception that narrowing down the type of constructions that express writers' intentions to clauses poses problems in multi-lingual applications: what is expressed in a clause in one language might more suitably be expressed in a different construction in another language. This case was made especially by Rösner and Stede (1992) and Carlson and Marc (2001) who consequently proposed extensions of the set of minimal units. The motivation for the inclusion of certain constructions (or the exclusion of others) is not always readily understandable. So, researchers comprising Carlson and Marcu (2001) and Lüngen et al. (2006) (cf. also paper 0.5) working in the RST paradigm, but likewise Wolf and Gibson (2005, 2006), opt for including certain pre-posed PPs like "On Monday," in the list of minimal units, whereas temporal adverbs that potentially convey the same information ("Yesterday,") are not included. One reason for this decision might be that those researchers are working on corpora based on journal texts, where temporal expressions preferably are of the variety they include in the list; so the decision to include one type of expression but not the other might be rather pragmatic than theory-driven.

Other work has been less intention-oriented than that of Grosz and Sidner and that in the RST paradigm, and consequently employed a different reasoning to select units as elementary discourse units. One line of research that also dates back to the mid-80s of the last century seeks to understand coherence in more general terms than tied up with linguistics. In (Hobbs, 1985) and work that can be seen in its tradition, like (Kehler, 2002), it is argued that coherence in text is by and large a product of the capability of rational agents to understand the world as being coherent. In this tradition, in its roots at least dating back to Hume (cf. (MacCormack & Calkins, 1913)), what is related by the rational mind are events or states of affairs. Consequently, what counts as a minimal unit in these accounts are expressions that can serve to convey states or events, or in short *eventualities*. A first class citizen here is the clause again, and once again with suitable restrictions excluding, e.g., restrictive relative clauses. For different reasons, Asher and Lascarides (2003)

and Reese et al. (2006, 2007) working in the Segmented Discourse Representation Theory (SDRT) paradigm too consider the expression of eventualities to be the decisive criterion for individuating minimal units.

As is well known from work in Montague grammar and elsewhere, it is all too easy to coerce the type of expressions to that of an eventuality. In fact, the paper by Jerry Hobbs in this collection (see 0.5) focuses on going below the clause level as minimal units. Hobbs there points out that he does think that even single words potentially express eventualities. It seems there is a thin line between raising types to that of an eventuality too easily and missing out on sub-clausal constituents that in fact are rhetorically interesting.

Yet another line of research different from the intention-oriented and the eventuality-based ones can be seen in processing-based accounts. One of the first candidates there, and also rooted in the mid-1980s is the work of Polanyi (1986, 1988); but also the work by Webber (2004) can be seen in that tradition. Polanyi's LDM most closely mirrors the incremental nature of text processing in that a discourse tree is built by adding sentences as elementary units to the existing representation of the text so far perceived. Sentences obviously are larger units than clauses since they potentially consist of multiple clauses (matrix, relative clauses, complement clauses). Both the LDM and Webber's D-LTAG suggest extensions to sentential syntax to model discourse structure, and thus there seems to be no need for sub-sentential segmentation.

Both LDM and D-LTAG — at least as concerns segmentation — thus seem to follow the opposite strategy than that pursued by Hobbs in his contribution in the present volume and reserve the rhetorical importance to larger units.

As a summary to the above approaches to segmentation, it seems that all accounts agree on a core set of units (main clauses that form sentences) that should be treated as elementary units, whereas there is large disagreement as to what else should be considered a unit in discourse.

0.3 Rhetorical relations

In Section 0.2 various views on how to split up discourse were reported. The present section is concerned with putting Humpty Dumpty together again: it is agreed among linguists that coherent discourse should be represented as a connected structure where each segment is connected to the rest by rhetorical relations. Islands in the representation of the analysis of a text are dispreferred and viewed as either a sign of incoherence of the discourse under analysis, faulty analysis itself or some lack in descriptive power in the inventory of rhetorical relations.

In what follows in this section, the accounts used to introduce segmentation strategies in the order chosen in Section 0.2 will be taken up in turn again and the rhetorical relations employed by those accounts will be sketched.

On the intention-oriented side, Grosz and Sidner (1986) employ a surprisingly small set of rhetorical relations. In their seminal paper they mention only two of them, one being *dominance* (the dominated discourse unit serves to achieve the goal of the super-ordinate) and the other *satisfaction precedence* (the preceding goal has to be achieved before the next can be achieved). The authors are aware of the fact that in, e.g., the work by Mann & Thompson (ultimately published in Mann and Thompson (1987a), but circulating in various grey versions beforehand) a much larger number of rhetorical relations are discussed. However, since for Grosz and Sidner primacy is on *intentions* and their relations to each other rather than on textual realizations of intentions, they can claim that dominance and satisfaction-precedence are sufficient for the description of rhetorical structure and the specific relations between textual units derivable from structure and intention content.

Grosz and Sidner (1986) explicitly set up their account for construction dialogues; given the goal orientation of that dialogue type, it seems that the inventory consisting of dominance and satisfaction-precedence suffices to describe the intentional structure of dialogues from that domain. This might be questioned, however, in a more general domain, where a putative task structure (if there is any) might be not as tightly bound to discourse structure. On the other hand, it has to be said that Grosz and Sidner (1986) do not *deny* the existence of more relations between discourse purposes. The claim, it seems, is just that the set suffices for the analysis of the given type of discourse.

As mentioned, Mann and Thompson (1987a, 1988b) posit a much larger set of discourse relations that should be used to describe the functional role of elementary units as recognized by the analyst: according to that classification, one unit might, e.g., *elaborate* on another, or units might form a *list*. There are two main divides within the class of relations: the first divide concerns the functional classification of relations: part of them are *subject-matter* relations (reporting about facts), another part is *presentational*, employed to influence the readers' stance towards the (main or local) discourse topic. Both of these types of relations can be realized in either of two ways (giving the second divide) — connecting a less important part of discourse (a *satellite*) to a more important one (called *nucleus*), or connecting nuclei to nuclei. The first type along the latter divide is called *mono-nuclear* relation (or nucleus-satellite relation), the second *multi-nuclear*.

One of the tests for nuclearity of discourse units is an elimination test: eliminating nuclei from a text tends to render it incoherent, while eliminating satellites

tends to leave coherence intact. This property of nuclei has led Marcu (1996) to posit the *nuclearity principle* for RST, claiming that spans of texts are connected by a relation iff their nuclei are. (A consequence of that principle for the representation of discourse structure will be discussed in 0.4.) According to Matthiessen and Thompson (1988), there is another indicator for nuclearity: they observed that there is a high correlation between the status of being a nucleus in a text and of being realized in a main clause just in case a rhetorical relation is present between two syntactically related clauses. Syntactically subordinate clauses tend to realize satellites in turn. As Matthiessen and Thompson (1988) warn, this is not a hard and fast rule, but rather a tendency, and counter examples abound. There even seem to be language specific discourse connectives that trigger an inversion in nuclearity, like the dutch connective *zodat* (*so that*) which in a majority of cases syntactically subordinates a nucleus to a satellite.

Bateman and Rondhuis (1994, 1997), and recently Stede (2008), systematically investigate rhetorical relations across different discourse theories and, for RST's nuclearity, propose not to tie the assignment of nuclearity to the presence of certain rhetorical relations (i.e., to drop the divide between mono-nuclear and multi-nuclear relations) and to view assignment of nuclearity as an effect of the presence of other factors. This seems to be in line with the findings by Asher and Vieu (2005) who claim something similar for an analogous divide among relations in SDRT. The insight that certain relations can be viewed as connecting nuclei to satellites *or* satellites to other satellites seems also to be the rationale behind the explosion of number of relations in the RST-flavor proposed by Carlson and Marcu (2001), where multi-nuclear versions of relations formerly categorized as mono-nuclear abound.

The discussion about the "right" relations for RST doesn't seem to be settled nor does it seem it has to be: Taboada and Mann (2006a) in their recent overview over developments in RST propose that researchers in the paradigm should tailor their own relations according to their specific needs for specific purposes.

The situation is different in SDRT, which, as mentioned, knows a similar divide as the nucleus-satellite distinction in RST. SDRT knows thorough axiomatizations of the discourse relations that are employed. These relations take the semantic representations of the minimal units and join them in either of two ways: by coordinating a unit to a preceding one, or by subordinating one unit to another. The nature of the relation involved (subordinating or coordinating) has influence on the possibilities where subsequent units can be attached: if the last relation involved was coordinating, then the constituent to which the last unit was related by it is blocked for attachment. If the last relation, on the other hand, was subordinating, then both the last unit and the one to which it was attached are available. These constraints on attachment points for new discourse units give rise to what is

called the *Right Frontier Constraint* (RFC), first postulated by (Polanyi, 1986). One effect of the RFC is to limit anaphoric accessibility: antecedents are said to be only available for (pronominal) anaphoric uptake if they occur in a unit that is on the right frontier.

Asher and Vieu (2005) re-examine the distinction between the two classes of relations and suggest that the question whether, e.g., a *cause*-relation is subordinating or coordinating depends in part on the information structure exhibited by the units that are related. Certain information structural configurations in the units can lead to anaphoric accessibility of discourse referents whereas truth semantically equivalent variants of that information structure makes them inaccessible. This fact can be accounted for if it is assumed that the information structure at least in part can change the way a unit is attached to preceding discourse.

SDRT draws another distinction between discourse relations that resembles the distinction between presentational and subject-matter relations in RST: many of the discourse relations are *content-level* relations which are similar to subject-matter relations, whereas other relations bear more resemblance to presentational relations, like *text-structuring*, *cognitive-level* discourse relations or *metatalk* relations. Interestingly, the so-called satisfaction scheme for veridical relations holds for some, but not all, relations of either variety. The latter scheme tells that two (representations of) discourse units connected by a relation are part of the interpretation of a discourse just in case the interpretations of the units are and the interpretation of the relation is. Whereas a veridicality criterion like that is to be expected for content-level relations, it is not so clear that a relation like *parallel* (a text-structuring relation) should have that property. None of the cognitive-level relations are veridical, though.

Just like SDRT and the account of Grosz and Sidner (1986), but unlike RST, the account of discourse relations given by Hobbs (1985); Hobbs et al. (1993) and, more or less based on it, Kehler (2002) is an attempt to give a principled way to define the ways units of discourse are combined to form larger units. Hobbs et al. (1993) distinguishes four classes of discourse relations: some that are inferred to hold because the units that are connected are about events in the world (like casual relations), others that relate what was said to an overall goal of the discourse, again others that relate a unit to the recipient's prior knowledge (e.g., *background*) and finally "expansion" relations (like *contrast*).

0.4 Structures and their properties

So, both what counts as minimal units and in which ways they can be combined by rhetorical relations are matters of dispute in discourse theory. Given this situation,

it can be expected that there is also no consensus on which structures discourse can be expected to have.

The expectation is confirmed by the literature. Whereas many researchers — e.g., Polanyi (1986, 1988, 2001), Grosz and Sidner (1986, 1998), Mann and Thompson (1987b, 1988a), Taboada and Mann (2006b) — assume that trees suffice to model the rhetorical structure of dialogue, there is an increasing number of theorists that doubt this assumption for a variety of reasons.

Prominent among the latter are Wolf and Gibson (2005, 2006) who argue for a much less constrained type of graphs for the description of rhetorical structure. The *chain graphs* they postulate as adequate for the description of rhetorical structure feature all kinds of violations of tree structure: they posit nodes with multiple parents, crossing edges, and in general graphs without a distinguished root node. Their strongest constraint on structures seems to be connectedness and acyclicity. These graphs are capable of describing all kinds of relations between elementary units; a closer look at their annotation manual and the set of relations they employ reveals that this seeming strength is a real weakness too: the set of relations Wolf and Gibson (2005, 2006) employ is a mixed bag, mostly taken from (Hobbs, 1985; Hobbs et al. 1993), but considerably modified and enriched with some relations from (Carlson & Marcu, 2001). The annotation manual requires analysts to annotate not only rhetorical relations used for combining minimal units, but also coreference relations and other cohesive devices that can be present within minimal units as well. Their first step of analysis, *grouping*, actually consists in connecting units that are related by cohesive links. Only after that very step rhetorical relations are applied to the units — alas not to the units *connected* by the first step. Thus, it is no wonder that crossing dependencies and nodes with multiple parents abound in the analyses presented in Wolf and Gibson (2005, 2006). Knott (2007) questions the statistics Wolf and Gibson (2005, 2006) perform on their data, claiming that the small percentage of tree violations that can be tied to the special relations introduced by Wolf and Gibson in their evaluation of their data is implausible. I don't think so: rather, Knott's critique seems to set in too late. The true reason for the high amount of tree violations does not lie in the special relations, but in the conflation of levels of analysis.

Another line of attack on tree structures as the adequate description for rhetorical structures can be found in (Danlos, 2004, 2008). Danlos compares the generative capacity of a comparatively unrestricted formalism (an extension of Mel'čuk's (Roberge, 1979) dependency syntax to discourse) with those of RST and SDRT. She derives all the structures that can be generated by either formalism and tries to find discourses that exhibit the respective structure. Her benchmark formalism generates directed acyclic graphs (DAGS), whereas RST and SDRT generate trees. According to Danlos analysis, RST undergenerates (is not complete), her

benchmark formalism overgenerates (is not correct), and SDRT is closest to being both complete and correct, with the exception being a few structures that can not be described as SDRT-trees. Danlos argument to my mind has two drawbacks, though it is admirably ingenious. First, it should be evaluated on corpus data instead of relying on constructed discourse for confirmation. This is mainly a precaution against an overreliance on intuitions, of course. The second point is a bit stronger: the reconstruction of RST mainly — based on (Marcu, 1996) and (Carlson & Marcu, 2001) — seems to contain too strong an interpretation of the nuclearity principle that leads to the assumption of graphs that are not in accordance with most other work in RST. So, rather than raising an argument against general RST assumptions, her attack is directed against a very idiosyncratic version of RST.

But both the account of Wolf and Gibson (2005, 2006) and of Danlos (2004, 2008) are under active discussion, and until there is conclusive evidence to the contrary, it has to be assumed that there are strong arguments against general tree-hood of discourse structure. A weaker warning comes from Webber (2001) and Lee et al. (2008): these authors caution that although most discourse can be modelled as trees, there might be certain cases where a departure from tree structures is required. So the warning would be to give up the general claim in favour of a rule of thumb with dened exceptions.

It seems that the question of how rich a structure has to be assumed for the description of discourse these days is more hottly debated than ever. The papers in the present volume will help to solve or focus in this debate by contributing insights in the fundamental questions that have to be answered.

0.5 About the papers

Jerry R. Hobbs: Clause-Internal Coherence

As was discussed in Section 0.2, there is no unanimity about the size or general characterization of elementary discourse units, just as there is no agreement on the definitions of relations between them. In his contribution to this volume, Hobbs extends his account, e.g., from (Hobbs, 1985), to cover coherence at a sub-clausal level.

Henk Zeevat: Optimal Interpretation for Rhetorical Relations

Zeevat argues that rhetorical relations can be reconstructed from general optimality theoretic (OT) assumptions. He gives a comprehensive introduction to OT, with special emphasis on the constraints *NEW, RELEVANCE, FAITH and PLAUSIBLE. He then continues to demonstrate how a range of rhetorical relations can be derived from a certain ordering of these constraints; most importantly, *NEW and

RELEVANCE tend to introduce rhetorical structure defaults, with PLAUSIBLE being a filter over the generated relations. This account of coherence relations is in marked contrast to accounts such as that of Hobbs (1985) or Asher and Lascarides (2003).

Ekatarina Jasinskaja: Modelling Discourse Relations by Topics and Implicatures: The Elaboration default

Jasinskaja argues in her paper for the position that discourse relations can be inferred by utilising underlying pragmatic principles such as *topic continuation* and *exhaustive interpretation* as defaults. In the absence of linguistic markers that make one more inclined to infer a different relation, she opts for ELABORATION as one of the default relations, since it best obeyes both principles, i.e., does not induce topic shifts and at the same time add information to the topic at hand.

Maja Bärenfänger, Harald Lüngen, Mirco Hilbert & Henning Lobin: The role of logical and generic document structure in discourse analysis

The authors of this contribution propose to add two descriptive levels to the local rhetorical analysis of discourse structure: the logical structure (like title, paragraph etc.) and the genre specific structure (introduction, method). Structure at these levels is usually not explicitly signalled, yet conventionalized, and can thus be used to guide (automatic) parsing of texts. The authors strive to clarify which cues and constraints can be observed at these levels of discourse and demonstrate their utility for automatic text processing.

Pascal Amsili & Claire Beyssade: Obligatory Presupposition in Discourse

Presupposition triggers have been considered obligatory under certain conditions by a variety of authors. One of the conditions that was deemed necessary in previous work was that the triggers are additive particles, like *too*. Amsili & Beyssade argue that this condition is not a necessary one, but that obligatoriness is the case for triggers that have no asserted content. (*Too* being but one of them.) They give a general explanation for the apparent sensitivity of this class of triggers to discourse relations and provide a formalization in terms of an SDRT update mechanism, building on Asher and Lascarides (1998).

Ann Copestake & Marina Terkourafi: Conventionalized speech act formulae — from corpus findings to formalization

Copestake & Terkourafi present an account to the semantics and pragmatics of conventionalized speech acts which renders the contribution of the illocutionary force as an addition to the compositional semantics of the utterance. They motivate their account with examples from a corpus of Cypriot Greek and formalize it within the framework of HPSG. They show how their account leaves the possibility for a literal interpretation of conventionalized speech act formulae open, thus opening the possibility to react to them in a variety of ways.

Philippe De Brabanter: Constraints on metalinguistic anaphora

De Brabanter reports the results of his research into a specific kind of referring expressions, where the referent itself is a linguistic object (like in " 'Boston' is disyllabic"). He draws a number of distinctions among those expressions, arguing that the class of metalinguistic anaphora referring back to expressions that themselves do have non-linguistic referents are especially interesting for a number of reasons.

Cyril Auran & Rudy Loock: Appositive Relative Clauses and their Prosodic Realization in Spoken Discourse: a Corpus Study of Phonetic Aspects in British English

In their contribution to the volume, Auran & Loock argue that differences in pragmatic functions fulfilled by appositive relative clauses are correlated to differences both in morphosyntactic and semantic characteristics and prosodic features. The latter mainly concern intonation, rhythm and intensity. The data they use are extracted from a corpus of spoken British English.

Bibliography

Asher, N. & Lascarides, A. (1998). The semantics and pragmatics of presupposition. *Journal of Semantics*, 15: 239–99.

Asher, N. & Lascarides, A. (2003). *Logics of Conversation*. Cambridge University Press.

Asher, N. & Vieu, L. (2005). Subordinating and coordinating discourse relations. *Lingua*, 115: 591–610.

Bateman, J. & Rondhuis, K.J. (1994). Coherence relations: analysis and specification. Technical Report R1.1.2: a,b, DANDELION Esprit Basic Research Project 6665.

Bateman, J. & Rondhuis, K.J. (1997). Coherence relations: towards a general specification. *Discourse Processes*, 24: 3–49.

Benz, A. & Kühnlein, P., editors (2008). *Constraints in Discourse*, volume 172 of *Pragmatics and Beyond new series*. John Benjamins.

Butterworth, B. (1975). Hesitations and semantic planning in speech. *Journal of Psycholinguistic Research*, 4: 75–87.

Carlson, L. & Marcu, D. (2001). Discourse tagging manual. Technical Report ISI-TR-545, ISI. http://www.isi.edu/marcu/discourse/tagging-ref-manual.pdf.

Chafe, W.L. (1980). *The Pear Stories*, volume 3 of *Advances in Discourse Processes*, chapter The Deployment of Consciousness in the Production of a Narrative, pages 9–50. Ablex.

Danlos, L. (2004). Discourse dependency structures as constrained dags. In Strube, M. & Sidner, C., editors, *Proceedings of the 5th SIGdial Workshop on Discourse and Dialogue*, pages 127–135.

Danlos, L. (2008). Strong generative capacity of rst, sdrt and discourse dependency *dags*. In Benz, A. & Kühnlein, P., editors, *Constraints in Discourse*, pages 69–96. John Benjamins.

Grosz, B. & Hirschberg, J. (1992). Some intonational characteristics of discourse structure. In *Proceedings of the International Conference on Spoken Language Processing*, pages 429–32.

Grosz, B.J. & Sidner, C. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3): 175–204.

Grosz, B.J. & Sidner, C. (1998). Lost Intuitions and Forgotten Intentions. In Walker, M.A., Joshi, A.K., & Prince, E.F., editors, *Centering Theory in Discourse*, pages 39–51. Clarendon Press.

Hirschberg, J. & Nakatani, C.H. (1996). A prosodic analysis of discourse segments in direction-giving monologues. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 286–93, Santa Cruz.

Hirschberg, J. & Pierrehumbert, J. (1986). The Intonational Structuring of Discourse. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pages 136–44.

Hirschberg, J. & Pierrehumbert, J. (1992). The intonational structuring of discourse. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pages 136–44. ACL.

Hobbs, J. (1985). On the coherence and structure of discourse. Technical Report 85-37, CSLI.

Hobbs, J., Stickel, M., Appelt, D., & Martin, P. (1993). Interpretation as abduction. Technical report, SRI International.

Kehler, A. (2002). *Coherence, Reference and the Theory of Grammar*. CSLI.

Kempen, G., editor (1987). *Natural Language Generation*. Number 135 in NATO Advanced Science Institutes—Applied Sciences. Martinus Nijhoff Publishers.

Knott, A. (2007). Review of "Coherence in Natural Language: Data Structures and Applications" by Florian Wolf & Edward Gibson. *Computational Linguistics*, 33(4): 591–5.

Lee, A., Prasad, R., Joshi, A., & Webber, B. (2008). Departures from tree structures in discourse: Shared arguments in the penn discourse treebank. In Benz, A., Kühnlein, P., & Stede, M., editors, *Proceedings of CID III*.

Lüngen, H., Puskás, C., Bärenfänger, M., Hilbert, M., & Lobin, H. (2006). Discourse Segmentation of German Written Texts. In *Advances in Natural Language Processing*, volume 4139 of *Lecture Notes in Computer Science*, pages 245–56. Springer.

MacCormack, T.J. & Calkins, M.W., editors (1913). *Hume, David: An enquiry concerning human understanding and selections from A treatise of human nature, volume 7 of Bibliotheca philosophorum*. Meiner.

Mann, W.C. & Thompson, S.A. (1987a). Rhetorical Structure Theory: A Theory of Text Organization. Technical Report RS-87-190, Information Sciences Institute, Los Angeles, CA.

Mann, W.C. & Thompson, S.A. (1987b). Rhetorical Structure Theory: Description and Construction of Text Structures. in: (Kempen, 1987). pp. 85–95.

Mann, W.C. & Thompson, S.A. (1988a). Dialogue Games: Towards a functional theory of text organization. *Text*, 8(3): 243–81.

Mann, W.C. & Thompson, S.A. (1988b). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8(3): 243–81.

Marcu, D. (1996). Building Up Rhetorical Structure Trees. In *The Proceedings of the Thirteenth National Conference on Artificial Intelligence*, pages 1069–74, Portland, Oregon.

Matthiessen, C. & Thompson, S.A. (1988). The structure of discourse and 'subordination'. In Haimann, J. & Thompson, S.A., editors, *Clause Combining in Grammar and Discourse*, volume 18 of *Typological Studies in Language*. John Benjamins Publishing Company.

Polanyi, L. (1986). The linguistic discourse model: Towards a formal theory of discourse structure. Techn. Report TR-6409, BBN Laboratories Inccap.

Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of Pragmatics*, 12: 601–38.

Polanyi, L. (2001). The Linguistic Structure of Discourse. In Schiffrin, D., Tannen, D., & Hamilton, H.E., editors, *Handbook of Discourse Analysis*. Blackwell.

Reese, B., Denis, P., Asher, N., Baldridge, J., & Hunter, J. (2006). Reference manual for the analysis and annotation of rhetorical structure. Technical report, Discor, Univ. of Texas, Austin.

Reese, B., Hunter, J., Asher, N., Denis, P., & Baldridge, J. (2007). Reference manual for the analysis and annotation of rhetorical structure (version 1.0). Technical report, Discor, Univ. of Texas, Austin.
http://comp.ling.utexas.edu/discor/manual.pdf.

Roberge, R.T., editor (1979). *Studies in Dependency Syntax — Igor A. Mel'čuk*. Karoma, Ann Arbor.

Rösner, D. & Stede, M. (1992). Customizing rst for the automatic production of technical manuals. In *Proceedings of the 6th International Workshop on Natural Language Generation*, pages 199–214, London, UK. Springer-Verlag.

Stede, M. (2008). RST revisited: disentagling nuclearity. In Fabricius-Hansen, C. & Ramm, W., editors, *'Subordination' versus 'coordination' in sentence and text*. John Benjamins.

Taboada, M. & Mann, W.C. (2006a). Rhetorical structure theory: looking back and moving ahead. *Discourse Studies*, 8(3): 423–59.

Taboada, M. & Mann, W.C. (2006b). Rhetorical Structure Theory: looking back and moving ahead. *Discourse Studies*, 8(3): 423–59.

Webber, B. (2004). D-LTAG: extending lexicalized TAG to discourse. *Cognitive Science*, 28: 751–79.

Webber, B.L. (2001). Computational Perspectives on Discourse and Dialogue. In Schiffrin, D., Tannen, D., & Hamilton, H., editors, *The Handbook of Discourse Analysis*. Blackwell.

Wolf, F. & Gibson, E. (2005). Representing Discourse Coherence: A Corpus-Based Study. *Computational Linguistics*, 31(2).

Wolf, F. & Gibson, E. (2006). *Coherence in Natural Language: Data Structures and Applications*. MIT Press.

Clause-internal coherence

Jerry R. Hobbs
University of Southern California

1. Introduction

About twenty years ago I was in a jogging group and was known in that group for my eagerness to take shortcuts. Our standard route included a stretch on a dirt road that became deep mud when it rained, but we had a longer detour on pavement for those days. One day after a rain, as we approached this stretch, I said,

(1) Let's take the muddy way.

Everyone laughed.

What's funny about this?

If I had said, "Let's take the short way," it would have been understood that the shortness used in my description also provided a motivation for taking that route. Since I used the word "muddy" instead, it seemed as if I wanted to run that way *because* it was muddy, as if I enjoyed sloshing through the mud.

Very often different parts of a clause are connected inferentially in this way even though there is no direct syntactic connection between them and even though there is no explicit lexical signal of the relation.

Among the relations we find are the kinds of coherence relations we find between successive clauses or larger stretches of discourse. In Hobbs (1985) and other papers, I present an account of discourse structure in which the coherence relations are seen as possible interpretations of the adjacency of segments in the discourse. For example, in

(2) Chris and Pat are studying. Don't bother them.

we not only need to figure out who "Chris", "Pat", and "them" refer to, what it would mean for them to study, what it would mean to bother them and to not bother them. We also need to figure out what is conveyed by the blank space between "." and "D". We need to explain what relation makes these two sentences a part of the same discourse.

The typical relations that are conveyed by adjacency of clauses, as argued in Hobbs (1985) include causal, interlocking change-of-state, similarity, and contrast

relations. Clauses describe events or situations, and these are some of the most common kinds of relations we express among events and situations. We see that the same set of relations can occur between different parts of a single clause, in a way that goes beyond the predicate-argument relations conveyed by syntax.

Consider another example:

(3) A jogger was hit by a car last night in Marina del Rey.

Our assumption is that the person was jogging when the accident occurred, and that somehow played a role in the accident. Contrast this with

(4) A professor was hit by a car last night in Marina del Rey.

Here we don't assume the professor's being a professor had anything to do with the accident. Professing is not an outdoor, on-street activity. In the sentence

(5) We should listen to the warnings of scientists.

there is an implicit causal relation between the listening and the scientists. It is *because* they are scientists, and therefore know what they are talking about, that we should listen to them. In

(6) Kids sometimes show great insight.

there is a violated expectation relation between the kids and insight. Normally children do not show great insight, and part of the message of this sentence is this violation of expected causality.

I will refer to events, situations, conditions, states, and so on as "eventualities". I have argued that nearly every morpheme in a sentence conveys some eventuality. For example, the sentence

(7) Sugar is sweet.

describes three eventualities, first, that some entity x is sugar; second, that x is sweet; and third, that *x's* sweetness holds at the present time.

Clause-internal coherence is the phenemenon of coherence relations holding between eventualities conveyed by morphemes in the clause, beyond the predicate-argument relations that are signalled explicitly by syntax.

Kronfeld (1989) discusses something like this problem under the heading of "conversationally relevant descriptions". In his account, in a sentence like

(8) The city with the world's largest Jewish community welcomes Israel's Prime Minister.

the reader recognizes that the definite description is not as brief as it could be, decides it must be conversationally relevant, and interprets it as containing an

implicit universal—any city with the world's largest Jewish community must welcome Israel's Prime Minister.

Kehler et al. (2008) use the expectation of clause-internal explanations to manipulate the resolution of attachment ambiguities in sentences whose main verb tends to invite an explanation. In a sentence that begins

(9) John detested the servant of the actress who ...

the reader expects an explanation of John's detesting the servant, and hence expects the relative clause to modify "the servant" and to provide the explanation, rather than expecting it to attach low to "the actress" as the ordinary syntactic default would indicate.[1]

In Section 2 of this Chapter I briefly describe the "Interpretation as Abduction" framework that provides a unified approach to a large number of linguistic phenomena. It contains a position on what counts as an adequate interpretation of a text, which clause-internal coherence challenges. In Section 3 I show how some cases of clause-internal coherence can be handled in the course of validating explicit linguistic signals or as examples of coreference. However, there are other examples, mostly involving similarity and violated expectation relations, that cannot be handled by these means. In Section 4 I consider a number of other examples from several different genres. Section 5 summarizes what we can learn from this investigation.

2. The Abduction framework

The key idea behind the "Interpretation as Abduction" framework (Hobbs et al. 1993) is that we interpret the world by coming up with the best explanation for the observables in our environment. The process of abduction is proving the thing to be explained deductively where possible, making assumptions where necessary, and deciding among alternative possible proofs by means of some measure of economy in proofs. When applied to interpreting language, the observable is the text itself. We need to explain why this text occurred. This divides into two subquestions: what information does the text conventionally convey (the informational perspective) and why did the speaker or writer want to convey this information (the intentional perspective).

Texts convey meanings via words (or morphemes) and via adjacency relations. Within individual sentences, the adjacency of words or larger stretches of text convey predicate-argument relations, and the recognition of syntactic structure

1. I am indebted to Andrew Kehler for drawing my attention to both his and Kronfeld's work.

is precisely the discovery of these relations. Thus, in the sentence "Pat works," the word "Pat" tells us that there is an entity x_1 named "Pat". The word "works" tells us that an entity x_2 works. The interpretation of the adjacency of the two words as conveying the Subject-Verb Phrase relation tells us that x_1 and x_2 are the same entity; Pat is the one who works. Thus, conceptually, the first step in the informational analysis of the sentences in a text is the discovery of the syntactic structure and the corresponding logical form.

One then tries to find the best abductive proof of the logical form of the sentences in the text, where the criteria for what is best include the following:

- Short proofs are better than long proofs.
- Salient axioms are better than nonsalient axioms.
- The fewer assumptions the better.
- A proof is better if it exploits implicit redundancies in the text.

The last of these in important in coreference resolution. One often gets a better proof by assuming that two things mentioned in different parts of a text are in fact the same, or are inferentially related. In the sentence

(10) I just bought a used car. It gets good gas mileage.

the most economical explanation happens if we assume "it" to be the same as "a used car". This is an example of *direct coreference*. The two references are to the same entity. In the sentence

(11) I just bought a used car. The tires are worn.

we get the best interpretation if we assume the tires are the tires of the used car, and not some random unrelated tires. The existence of the car proves the existence of the tires. This is an example of *indirect coreference*.

Sometimes we get the best interpretation by recognizing only a partial match. In

(12) I talked to my mechanic. She said the engine is in good shape.

both "mechanic" and "she" refer to persons, but both carry properties that are not contained in the other—occupation and gender. Assuming the persons implicit in each are the same gives us the best proof of the logical form and resolves this direct coreference by the partial overlap in their meaning. In

(13) My tires are worn, but the engine is okay.

neither the tires or the engine implies the existence of the other, but if we assume a common core of "car", then the existence of both can be inferred. Again there is a partial overlap in meaning, and recognizing this allows us to resolve the indirect coreference.

Because abduction finds a proof of the logical form by backchaining, the interpretation is more specific than the explicit content of the sentence. The disambiguation of word sense ambiguities is one example of this. In the sentence

(14) The plane taxied to the terminal.

the plane could be an airplane or a wood smoother, taxiing could be a plane moving on the ground or someone riding in a taxi, and a teminal could be an airport terminal or a computer terminal. But by identifying the common core of the first reading of each, we get a much more economical interpretation of the whole sentence, thereby disambiguating the word sense ambiguities as a by-product. Where we refer to "explicit signals" in Sections 3 and 4, we are talking about lexical items or punctuation that convey very general meanings but get more specific meanings as a result of abductive interpretation.

I said above that within a sentence, adjacency conveys a predicate-argument relation; in fact, we can define sentences as that region of a text in which adjacency is conventionally interpreted as conveying a predicate-argument relation. Beyond the sentence, adjacency also conveys information, but the constraints are much looser. Adjacency says that the two clauses or larger stretches of discourse are somehow related. Clauses and larger stretches of discourse generally describe states and events, or eventualities. Thus, the relations most frequently conveyed by adjacency will be the kinds of relations that normally occur in the world among states and events. Overwhelmingly, these tend to be positive or negative versions of interlocking change of state, or what I have called the occasion relation; enablement, causality, and implication, and the negation of these in the violated expectation relation; the figure-ground relation; similarity or parallelism, and its negation, contrast; and a limiting case of similarity, the elaboration relation, or event coreference.

The occasion relation occurs when the first segment of a pair describes the initial state or a change into the initial state of a change of state described by the second segment.

(15) I flew to Paris. I traveled through France, Spain, and Italy.

The final state of flying to Paris is being in Paris, which is the initial state of the traveling.

The causal relation occurs when one of the segments describes a cause and the other an effect.

(16) I went to Europe. I needed a vacation.

Needing a vacation is the cause of going to Europe.

The violated expectation relation occurs when an expected causal or implicational relation does not hold. It is normally, but not always, signalled explicitly, e.g., by "but".

(17) I needed a vacation, but I spent the summer working.

Normally, needing a vacation causes one to not work.

The similarity or parallelism relation occurs when the same predication is made of entities that are similar in that, roughly, they share the same properties. A more precise definition of similarity is given in Hobbs and Kehler (1997).

(18) I flew to Paris. My wife flew to London.

The predication is flying. I and my wife are similar in that we are both people. Paris and London are similar in that both are major European capitals.

The contrast relation occurs when a property is predicated of one of two similar entities and its negation is predicated of the other.

(19) I took a taxi to the Louvre. My wife walked.

Riding in a vehicle implies that one did not walk, so a property and its negation are asserted of the similar entities, me and my wife.

The elaboration relation occurs when the two segments describe the same eventuality, perhaps from a different perspective.

(20) The Louvre is France's biggest tourist site. Millions go there every year.

Being the biggest tourist site and having millions visit are two descriptions of the same situation. It should be pointed out that this elaboration relation is not the same as the elaboration relation in Rhetorical Structure Theory (Mann & Thompson, 1986). Their elaboration relation is no more than coreference; two successive segments of text predicate properties of the same entity. In this elaboration relation, the eventuality whose existence is asserted or claimed by the two segments have to be the same. Thus, the pair

(21) France's biggest tourist site is the Louvre. It was built in 1202.

would be an elaboration in RST, but it would not be in my account. The Louvre's being the biggest tourist site and its being built in 1202 are not the same eventuality.

The "Interpretation as Abduction" framework can be seen as an answer to the question, "What inferences should we draw from a text?" The answer is those that contribute to the best explanation of the fact that the text occurred, which in part is the best proof of its logical form. The two aspects of this most relevant to this paper are that direct and indirect coreference resolution and the discovery of specific interpretation of general predicates both fall out of the interpretation as a by-product of the process of finding the most economical proof.

3. Classes of clause-internal coherence

3.1 The data

In the paper as a whole, I will examine texts from four diverse sources to see what instances of clause-internal coherence we find and how they can be recognized in the abduction framework. The four sources are

- A *Science* article on AIDS.
- An article from the business section of the San Jose *Mercury-News.*
- The first paragraph of Carson McCullers' *Ballad of the Sad Cafe.*
- Shakespeare's 64th sonnet.

In this section, I will examine three classes of cases: clause-internal coherence indicated by explicit signals; clause-internal coherence that falls out in the same way as coreference falls out in interpretation by abduction; and a residue of cases that remain problematic for the abduction framework. The three cases are illustrated in the following sentence opening the business news article:[2]

(22) In a stunning reversal for one of Silicon Valley's fastest growing companies, Media Vision Technology Inc. said Thursday it will report a sharp decline in sales and a "substantial loss" in the quarter ending March 31—a jolt that cut its stock price in half.

3.2 Clause-internal coherence from explicit signals

Recognizing clause-*external* coherence is a matter of interpreting adjacency. That is, the coherence relation is the best explanation of why the two discourse segments are next to each other. But many of these relations are in addition explicitly signalled by a conjunction or discourse adverbial. Similarly, many instances of clause-*internal* coherence are also explicitly signalled.

In Example (22), there is a similarity between "sharp" and "substantial", and a similarity between "decline" and "loss". But the phrases "a sharp decline in sales" and "a 'substantial loss' " are conjoined by "and".

The word "and" rarely just means logical conjunction. It has two principal specializations: "and then" and "and similarly". The latter is probably more common. Specialization of the information in the logical form of a text is precisely what abduction does. Finding this specialization of "and" is equivalent to discovering the similarity. The similarity rests on the fact that a decline in sales and a

2. From the San Jose *Mercury News,* March 25, 1994, p. 12E.

loss (in profits) are both drops in postive measures for a business, and "sharp" and "substantial" both indicate the high region of the scales measuring these drops.

We may not normally think of "and" as conveying much information. But here it is the explicit signal that drives the recognition of the clause-internal coherence relation of similarity.

3.3 Intra-clausal coherence as coreference

There is an elaboration relation between the word "stunning" and the phrase "a jolt that cut its stock price in half". We can view the condition of being stunning and the condition of being a jolt as two descriptions of the same situation. That is, by unpacking their meaning into a common core of something like "a sudden and surprising event", and assuming these are the same, as a way of getting the most economical abductive proof, we thereby recognize the elaboration relation and in a sense see the two conditions, the jolt and the "stunning-ness", as coreferential.

There is a violated expectation relation between "fastest growing" and "reversal"; normally, fast growth leads to more growth. Suppose for the moment we could recognize this. Then that violated expectation proposition becomes part of the interpretation of the text and is itself available as a source of inferences. Violated expectations cause surprise, and this relation is thus a partial explanation of the existence of the condition of being stunning and the condition of being a jolt. This is a kind of indirect coreference; the violated expectation exists, and therefore the "stunning-ness" exists.

In both of these cases there is an implicit redundancy in the situations described by different words or phrases in the text. We get a better abductive proof by assuming these situations are identical. Where the result is the resolution of direct coreference, we have discovered an elaboration relation. Where the result is the resolution of indirect coreference, we have discovered either an implication relation, or something more specific encoded in the axiom we use, such as causality or enablement.

3.4 Problematic residue

Let us return to the violated expectaton relation between "fastest growing" and "reversal". If we saw the text

(23) Media Vision Technology Inc. had been growing fast. There was a reversal today.

we would have to interpret the adjacency of the two sentences by finding a coherence relation between their claims, i.e., the fast growth and the reversal. This would drive the recognition of a clause-external violated expectation relation.

Within clause boundaries, however, adjacency only conveys predicate-argument relations, and in any case, the phrases are not adjacent.

The explicit signal that comes closest to relating the reversal and the fast growth is the preposition "for". But "for" expresses a relation between the reversal and the company, which just happened to have been growing fast. There is no direct syntactic relation between the reversal and the growth, so we cannot hope to discover the coherence relation by means of an explicit signal.

If things of type X cause or imply things of type Y, then we can normally at least partially infer one from the other. But the violated expectation relation happens precisely when that expected causal or implicational relation does *not* hold. So we can't expect to infer the reversal from the fast growth or vice versa, as a kind of indirect coreference.

It is true that a reversal requires some kind of directed motion to be reversed, and growth is just such a motion. Assuming the reversal allows us to partially prove the existence of the growth. However, this would only give us the occasion relation. The growth sets up the occasion for the reversal to happen. It does not yield the recognition of the violated expectation relation.

This example illustrates the problematic residue of cases, for which it is hard to see how the discovery of the clause-internal coherence relation would happen in the abductive framework.

In the next section we analyze a number of other examples of clause-internal coherence in our sample texts, and classify them into one of these three classes: explicit signal, coreference, or residue.

4. Further examples of clause-internal coherence

4.1 *Science* article

In the *Science* article on AIDS we find the sentence,[3]

(24) For a short but variable period—a few weeks to a few months— after an individual is infected with HIV-1, virus is typically found in the blood (viremia), and high levels of virus replication can be observed.

Explicit Signal: There is a contrast between "short" and "variable". It is explicitly signalled by "but". In fact, this is an example of a common pattern that might be

3. From p. 964 of "Antigenic Diversity Thresholds and the Development of AIDS", by Martin A. Nowak, Roy M. Anderson, Angela R. McLean, Tom F. W. Wolfs, Jaap Goudsmit, and Robert M. May, *Science,* November 15, 1991, pp. 963–969.

called a "Second-Order Refinement". The pattern is "X but not completely X". The first part makes a first-order approximation to the intended state, and the second part makes corrections at a finer granularity. Other examples are "Pat is tall, but he stoops a lot," and "Chris is an A student, but sometimes she makes big mistakes." This pattern is the source of the "Yes, but ..." construction. It expresses a contrast, but it blends into the Violated Expectation coherence relation since the first part defeasibly implies the state without the correction.

The phrase "a few weeks to a few months" is an elaboration of "short" (and of "variable"). There are often explicit signals of coherence relations that we might overlook in the analysis of a clause. In this case the explicit signal is the dash, "–". This normally signals an appositive construction and hence an elaboration. The appositive is providing a different description of the same situation. In this case, we reason that weeks and months are both periods. The word "to" indicates a range of values, as does "variable". The word "short" applied to periods and the word "few" applied to weeks and months describe the same situation. Hence, the elaboration relation signalled by the dash is validated.

There is a similarity relation between "virus is typically found in the blood" and "high levels of virus replication can be observed"; this is signalled by the word "and". As noted above, similarity is one of the most likely specializations of "and". Here the similarity is based on the fact that finding and observing are both acts of perception, the implicit agents of the perception in both cases are medical personnel, and what is found is the presence and activity of virus. In both cases, what is asserted is a perceiving of diagnostic properties of the virus.

Finally, there is an enablement relation between the infection, on the one hand, and the finding of HIV-1 in the blood and the observation of high levels of virus replication, on the other. Infection, precisely, is the physical transfer of the virus from outside the body to inside the body and its establishment there as a replicating entity. Its presence in the body enables the finding, and its replicating enables the observation of the replicating. Enablement implies a temporal relation, and this relation can be expressed by the word "after". In the abduction framework, one tries to prove the "after" relation, along with the rest of the logical form, and the enablement relation between the infection and the diagnosis is what proves it. Enablement is the more specific interpretation of "after" found by the abduction process.

The explicit signals in this example are "but", "—", "and", and "after".

Later in the *Science* article on AIDS we find the sentence,[4]

4. P. 694.

(25) Antibodies then appear in blood serum, after which it becomes difficult to isolate the virus; viral antigens are often undetectable during the long but variable asymptomatic or incubation period between primary HIV-1 infection and the occurrence of AIDS.

Explicit Signal: Three clause-internal coherence relations are explicitly signalled.

There is a causal relation between the appearance of antibodies in the blood serum and the difficulty of isolating the virus; it is the job of the antibodies to destroy the invading virus. This causal relation is explicitly signalled by the word "after", so the analysis is similar to Example (24) above.

The clause "viral antigens are often undetectable ..." is an elaboration on the difficulty. Here the explicit signal is the semicolon. A semicolon ";" can represent a number of possible relations, but elaboration is one of them. Recognizing the elaboration rests on lining up the implicit negative in "difficult" with the "un-" of "undetectable", and recognizing that what is negated in the two cases, the isolating and the detectability, are the same in that both are a matter of discovering the virus. Strictly speaking, this is an *inter*-clausal coherence relation, and would be recognized in the process of explaining the adjacency of the two clauses, regardless of the punctuation.

The words "asymptomatic" and "incubation" are in a kind of contrast relation; in fact, they stand in a function-structure relation, with "incubation" describing what is going on structurally or internally, and "asymptomatic" describing what is visible to the exterior in terms of the functioning of the entity. This contrast is explicitly signalled by the word "or", which here is used in a kind of speech act sense. It could be paraphrased as "the period which could be called 'asymptomatic' or could be called 'incubation'". That is, "asymptomatic" and "incubation" are embedded in a kind of metonymic operator— "could be called …". The contrast rests on the fact that "asymptomatic" means that nothing is happening, and "incubation" means that something is about to happen.

The explicit signals in this example are "after", ";", and "or".

Coreference: There is an elaboration relation between the propositions conveyed by "undetectable" and "asymptomatic". This can be recognized in the same way that coreference resolution happens, as a by-product of abduction. Symptoms are what enable someone to detect something. The words "undetectable" and "asymptomatic" both mean that there are no outward signs of the condition. They describe the same situation The lowest-cost proof of the logical form of the sentence will be one that assumes this absence is the same in both cases. The identification of these two situations is equivalent to recognizing the elaboration relation.

There is an occasion relation between the situations conveyed by the word "incubation" and the phrase "the occurrence of AIDS". The incubation period of

a disease is the time between infection with the vector and the occurrence of the disease. If the incubation happens, the occurrence will happen. Thus, the occurrence need not be assumed; it can be proved after assuming or proving the incubation, yielding a lower-cost proof and explicitating the inferential relation between them. Because the incubation of a disease is a change of state into the state in which the disease occurs, the relation between them is the occasion relation.

In the first of these two cases, we can say that the undetectability and the asymptomaticity are directly coreferential. In the second case we can say the incubation and the occurrence of the disease are indirectly coreferential. The existence of the first implies the existence of the second.

The next sentence in the article on AIDS is the following:

(26) The incubation period is characterized by low viral replication (interspersed with minor and short-lived upsurges of viremia in some patients), and by constant or slowly decreasing numbers of $CD4^+$ cells.

Explicit Signal: The two issues in this sentence are the level of viral replication and time. These are similar in that they are both parameters. They are contrasting in that time is the independent parameter and the level is the dependent parameter. Recognizing that they are both parameters and that the level depends on time requires us to interpret the word "characterized" correctly. If X characterizes Y, then there is a functional relation from Y to X. We can think of the statement "Red hair characterizes Irishmen" as positing a function from people to hair color that maps Irishmen into red hair. Here, different levels characterize different periods of time.

The word "with" also conveys that functional relation. Its first argument is "interspersed", which indicates a temporal aggregate, a set of temporal intervals. For these intervals to be "with" upsurges means that there is a mapping from each of the elements of the aggregate to the upsurge that characterizes it.

If we had in our knowledge base an axiom that said that elements of the domain of a function are independent parameters while elements in the range are dependent parameters, then we could in principle recognize the contrast relation.

Residue: The word "period" refers to a temporal interval. The word "interspersed" describes an aggregate of intervals all contained within a longer interval. The phrase "short-lived" describes the length of an interval. The phrase "slowly decreasing" says something about the length of the interval occupied by the decreasing event. Thus, all of the eventualities described by these words and phrases are in a similarity relation, by virtue of their reference to a temporal interval.

The word "low" describes the level of viral activity. The word "upsurges" describes episodes in which the level is higher than usual. The word "minor" moderates that, in a kind of Second-Order Correction contrastive relation with "upsurges", so it also refers to the level of viral activity. The phrase "numbers of

$CD4^+$ cells" describes a measure of the level of viral activity, and consequently the words "constant" and "decreasing" describe such levels. Thus, all of the eventualities described by these words and phrases are in a similarity relaiton, by virtue of their reference to level of viral activity.

These clause-internal similarity relations exist. But it is not clear how they would be recognized in the abduction framework. We can axiomatize similarity in a way that allows it to be validated when it is explicitly signalled. But here it is not. For example, in "interspersed with short-lived upsurges", the fact that "interspersed" and "short-lived" both make reference to time is, in some sense, accidental. We could have as felicitously said "interspersed by unexpected upsurges", where there is no similarity relation.

We cannot recognize the similarity as a kind of direct anaphora. They are not the *same* intervals or the *same* levels, so we would not want to identify them, on the way to finding the lowest-cost proof.

To recognize them as a kind of *indirect* anaphora, we would need an axiom that said for any entity, there is a similar, nonidentical entity. This axiom seems to be much too powerful.

Nevertheless, people are very very good at spotting similarities wherever they occur.

4.2 Business news

In the article from the San Jose *Mercury-News* business section, the two sentences immediately after Example (22) are the following:

(27) Media Vision plummeted to 11, down 10 1/2 in frantic NASDAQ trading as 14.2 million shares were traded, more than 25 times normal volume. Thursday's decline continues a precipitous two-month slide from a peak of 45 1/4 Jan. 20 that has wiped out $480 million in market value.

Explicit Signal: There is an occasion relation between the two-month slide and Thursday's decline. The slide is a change of state into a final state that was the initial state in the decline. But this is explicitly signalled by the verb "continues".

In addition, when X continues Y, the implication is that X and Y are the same sort of eventuality. That is, X and Y are similar. In this case, they are similar because both are downward motion.

Coreference: The word "plummeted" indicates a rapid movement downward. The word "down" of course indicates a downward direction. The word "decline" also refers to a downward movement, as does slide. Money is a metaphorically vertical scale, so that decreasing a measure on that scale is a downward movement; the phrase "wiped out $480 million" thus indicates a downward movement. All of the eventualities conveyed by these words and phrases at least stand in a similarity relation be virtue of the downward movement they all indicate.

But in fact there are coreference relations involved here. The word "down" describes the same as the downward motion implicit in the plummeting. Thursday's decline is the same as the plummeting event. Thus, assuming a downward motion at least partially accounts for all three eventualities. Moreover, the wiping out and slide are both descriptions of the same event, and can similarly be recognized by identifying the two downward motions. The statement that the decline continues the slide should block coreference resolution between the two sets. Thus, we have an elaboration relation among the eventualities in each of the sets. Because of the word "continues", we have an occasion and a similarity relation between the first set and the second.

There is also a similarity among the various indicators of the intensity of that decline, including "plummeted", "frantic", "14.2 million shares", "more than 25 times normal volume", "precipitous", "wiped out", and "$480 million". Of course, for the numeric indicators to be recognized as signals of intensity, one has to know the normal range of values. There seem to be three situations being described here. For a process to be frantic is for its subevents to occur in rapid succession driven by fear in the agents. To plummet is to drop rapidly, so there is an identity in the impicit rapidity of the plummeting and the "frantic-ness". But the two situations themselves are not identical. The plummeting does not necessarily have to be driven by fear and the "frantic-ness" does not necessarily involve downward motion. The identity of the trading of 14.2 million shares and the volume of 25 times normal is indicated by the fact that the latter phrase acts as an appositive on the former phrase, once we coerce the latter from the verb "traded" to its subject (cf. Hobbs, 2001). Recognizing the other identities of eventualities depends on fairly complex reasoning about rates. For 25 times normal trading to occur in one day must mean that the trading was rapid, linking with the rapidity implicit in "plummeted" and "frantic".

The eventuality conveyed by "precipitous" is downward motion with a steep slope. This situation is described again in the reference to wiping out $480 million. To recognize this we must know that $480 million is a lot of money and is a loss that occurred in a two-month period, and thus is a quantitative measure of a slope.

The precipitousness of the slide is distinct from the rapidity of the plummeting simply because we can establish that the slide and the plummeting are not the same.

4.3 The novelette

The first sentence of the Carson McCullers novelette is as follows:[5]

5. Carson McCullers, "Ballad of the Sad Cafe", Houghton-Mifflin Co., Boston, MA, 1943.

(28) The town itself is dreary; not much is there except the cotton mill, the two-room houses where the workers live, a few peach trees, a church with two colored windows, and a miserable main street only 100 yards long.

Coreference: There is a causal or implicational relation between the cotton mill and the workers. A cotton mill is a factory and factories have workers, so we can see this inferential relation as a kind of indirect coreference.

There is also an implicational relation between the houses and the people who live there. Houses are where people live, and workers are people. There is thus a partial proof from "houses" of the existence of the workers, and a full proof of the existence of the living.

Explicit Signal: There is an exemplification, or more properly, "bad whole - bad part" relation, between the "dreary" town and the "miserable main street". There is also a similarity relation among the various indicators of quantity in the sentence—they are all small—including "not much", "two-room (houses)", "few (peach trees)", "two (colored windows)", and "only (100 yards long)".

It is possible that all of these should fall out from a recognition of the coherence relation between the two clauses, signalled by a semicolon, and an interpretation of "and" as "and similarly". The first clause describes the town as a whole. The second clause elaborates on this by describing its various parts—its factory, houses, trees, church, and street. These are all similar in that they *are* all parts of a town. Furthermore, all of these items except the cotton mill have quantitative descriptors. This adds to the similarity. But it is more important that the quantities are all small. This is explicit in "a few peach trees" and "only 100 yards long". To recognize "two-room houses" and "a church with two colored windows" as indicating a small quantity we need to know the normal range of quantities. The interpretation of "and" as "and similarly" spreads across the entire conjoined noun phrase object of the preposition "except", and the similarity of these items is established since all are small examples of things found in town. The preposition "except" conveys a Second-Order Correction. The phrase "not much" gives us a baseline pretty close to nothing, and the "except" phrase gives us a more detailed accounting of how it differs from nothing.

Not all small quantities are bad. If the town had few crimes and few tornadoes, that would be good. But we know that small house size, small church size, and a small number of trees is usually not very good. Thus, the quantitative properties all share a badness property with the dreariness of the town. We thus get an elaboration relation between the first clause—the town is bad—and the second—its parts are bad. The adjective "miserable" modifying "main street" fits into this pattern.

Thus, the recognition of the clause-internal similarity relations is driven by the interpretation of "and" and by the recognition of the elaboration relation between the two clauses.

The second sentence in the novelette is

(29) On Saturday the tenants from near-by farms come in for a day of talk and trade.

Coreference: There is an implicational relation between "Saturday" and "day". Saturday is a day, and the lowest cost proof results by assuming they are the same day.

There is also an implicational relation from the farms to the tenants. Farms have farmers who work them, and a tenant is a farmer who does not own the farm, so there is a partial proof of the existence of the tenant if we assume the existence of the farm.

The fact that the farms are "near-by" enables the tenants to come in, and this coming in in turn enables the talk and trade. Qualitative scalar concepts like "near-by" are generally associated with functional properties. If two points are literally or metaphorically near each other, then it is easy to traverse the distance between them. Thus, the nearness enables the coming in. Social interactions require the participants to be in the same location (which in the electronic age can be defined by a network of connectivities). So the coming into town enables the talk and trade. If we assume the talk and trade occurred, then we can assume the enabling identity of location occurred, so we can assume the coming together in that place occurred. If that is so, an enabling nearness to the starting points must hold as well. Thus, axioms with the structure "If X occurs, then Y defeasibly enabled it" allow us to infer the existence of the coming in and the nearness from the assumption of the talk and trade. These enabling relations are thus a kind of indirect coreference resolution.

4.4 Shakespeare's sonnet

The next four examples come from Shakespeare's 64th sonnet, and because it is sometimes hard to understand lines from it in isolation, the entire sonnet is presented here.

> When I have seen by Time's fell hand defaced
> The rich, proud cost of outworn buried age,
> When sometime lofty towers I see down-rased
> And brass eternal slave to mortal rage;
>
> When I have seen the hungry ocean gain
> Advantage on the kingdom of the shore,
> And the firm soil win of the wat'ry main,
> Increasing store with loss and loss with store;

When I have seen such interchange of state,
Or state itself confounded to decay,
Ruin hath taught me thus to ruminate,
That Time will come and take my love away.

This thought is as a death, which cannot choose
But weep to have that which it fears to lose.

In this sonnet, the poet observes that time destroys everything, and that time will eventually destroy his love. It is a poem about entropy.

The first example from the sonnet is the first quatrain:

(30) When I have seen by Time's fell hand defaced
The rich, proud cost of outworn buried age,
When sometime lofty towers I see down-rased
And brass eternal slave to mortal rage;

Coreference: Defacing is one instance of something being fell. The two words overlap in meaning; there are other ways of being fell, and defacement describes the process as well as merely the result. But this common core of meaning in the two words represents the same underlying situation. Assume this, and there is a partial proof of both the condition of being fell and the defacing event, thereby capturing the elaboration relation between them.

There is a causal relation between "Time" and the defacing. The passage of time causes things to be no longer intact. This is one aspect of being defaced, so the defacement can be seen partially as a indirect coreference from Time. The passage of time exists, so possibly a defacement does too.

There is also a causal relation between the "time" of "sometime" and "down-rased". This can be analyzed in the same way as the relation between "Time" and "defaced", as indirect coreference.

Residue: The whole quatrain is built on the contrast between valuable and intact things—"rich, proud cost", "lofty towers", and "brass eternal"— and the condition of being broken—"fell", "defaced", "outworn buried", "down-rased", and "slave to mortal rage". The general pattern is "Valuable and intact things break." We can recognize an Occasion relation from the intactness to the breaking, since breaking is a change of state whose initial state is intactness. But as in our analysis of Example (22), this does not give us the contrast relation.

The second quatrain is

(31) When I have seen the hungry ocean gain
Advantage on the kingdom of the shore,
And the firm soil win of the wat'ry main,
Increasing store with loss and loss with store;

Coreference: There is a metaphorical causal relation between "hungry" and "gain" (or "gain advantage")—the ocean's hunger causes it to eat the land. Eating is consuming, and consuming causes a gain. Thus, the causal relation can be discovered as a kind of indirect coreference resolution. The existence of the hunger causes the existence of the gain.

Residue: There is a clause-internal contrast betwen "firm" and "wat'ry". To be firm is to be solid; to be watery is to be not solid. But neither the firmness nor the watery-ness imply each other, so a relation cannot be discovered by coreference resolution. There is no explicit signal of the contrast. The word "and" indicates a parallelism between the first two lines and the third line, but the parallelism rests on one domain consuming the other. It does not go down to the level of the descriptors of the domains. That is, the hunger of the ocean bears no relation to the firmness of the soil.

The third quartrain is

(32) When I have seen such interchange of state,
Or state itself confounded to decay,
Ruin hath taught me thus to ruminate,
That Time will come and take my love away.

Coreference: There is a clause-internal contrast between the second occurrence of "state" and the word "decay". The sense of "state" here is "majesty, royalty, or splendor", but the poem works just as well if we take "state" to mean the more modern sense of a politically organized body of people under a single government. In either case we have highly structured entity. Decay is the loss of internal structure in a structured entity, and confounding is a cognitive version of the process of losing orderly structure. The higher level of structure in the start state of a decay process is implicit in the word "state". This is thus an example of partial indirect coreference.

There is an enablement relation between "come" and "take . . . away". If we assume the existence of the taking away event, it must have had its locational enabling conditions hold, and that is the final state of the coming. Thus, we can see the recognition of this enablement relation as a partial solution to an indirect coreference problem.

The final couplet is

(33) This thought is as a death, which cannot choose
But weep to have that which it fears to lose.

Residue: There is a clause-internal violation of expected causality between "weep" and "have"—normally it is a loss that would be wept, not a possession. We cannot

discover this as a variety of coreference, because neither the having nor the weeping imply the other. The relation between the two eventualities is, in more modern terms, "weep at having", where "at" conveys causality. But there is no rule that says that having what we want causes us to weep. Quite the contrary. So there is no mechanism in the abduction framework that would force the discovery of this clause-internal coherence relation.

5. Summary

The dense clause-internal coherence structure we have seen might be expected in poetry, and perhaps in the novelette as well, but the examples from the business news and the *Science* article indicate that the phenomenon is quite pervasive in all *written* discourse, at least. (Examples are harder to find in conversational data.)

The interpretation of many of these examples falls out from ordinary interpretation by abduction. The lowest cost explanation of the text as a whole contains the coherence relations within it. There are two mechanisms by which this can occur.

The first mechanism is the process of finding more specific meanings in context for words and other explicit signals than they would convey in isolation. Many of these signals might often escape our notice as requiring interpretation, such as words like "and" and punctuation like dashes and semicolons. The situation is analogous to when an inter-clausal coherence relation is signalled not only by adjacency but also by a conjunction or a clause-level discourse adverbial. In both cases, the abduction framework dictates that we prove abductively the very general meaning conveyed by the signal, and the specific coherence relation falls out of that proof.

The second mechanism resembles coreference resolution. The lowest cost abductive proof results if we assume that entities or eventualities described or implied by different parts of a text are in fact identical. Where they are *described*, we have thereby recognized an elaboration relation, analogous to *direct* coreference. Where one or more are merely *implied*, we have thereby recognized at least an implicational relation, analogous to *indirect* coreference, and if there is causality or enablement encoded in the axioms we use, we have thereby recognized those relations as well. Example (3) is just such a case, although a somewhat complex one. If a car hits someone, they must be in the same place. The location of the victim enables the accident. Jogging is usually done outside, and often joggers are in the street, which is where cars usually are. The existence of the jogger partially implies being located in the street, which enables the accident. This is thus a case of partial indirect coreference resolution in which enablement is part of the supporting abductive proof.

However, problematic cases remain. When the relation is one of violated expectation, we cannot expect the same sort of coreference based on an inferential or causal relation, because that is precisely what is violated. When the relation is one of similarity, there is no direct coreference to discover, since the eventualities are not identical. There is no indirect coreference to discover, since the existence of an eventuality does not imply the existence of similar eventualities. These cases constitute a challenge to the "Interpretation as Abduction" picture of what counts as an interpretation.

References

Hobbs, Jerry R. 1985. "On the coherence and structure of discourse." Report No. CSLI-85–37, Center for the Study of Language and Information, Stanford University.

Hobbs, Jerry R. 2001. "Syntax and metonymy." In *The Language of Word Meaning*, Pierrette Bouillon & Federica Busa (eds), 290–311. Cambridge, United Kingdom: Cambridge University Press.

Hobbs, Jerry R., & Kehler, Andrew. 1997. "A theory of parallelism and the case of VP ellipsis." *Proceedings*, 35th Annual Meeting of the Association for Computational Linguistics, 394–401. Madrid, Spain, July 1997.

Hobbs, Jerry R., Stickel, Mark, Appelt, Douglas, & Martin, Paul. 1993. "Interpretation as abduction." *Artificial Intelligence*, 63 (1–2): 69–142.

Kehler, Andrew, Kertz, Laura, Rohde, Hannah, & Elman, Jeffrey L. 2008. "Coherence and coreference revisited." *Journal of Semantics*, 25: 1–44.

Kronfeld, Amichai. 1989. "Conversationally relevant descriptions." *Proceedings*, 27th Annual Meeting of the Association for Computational Linguistics, 60–67. Vancouver, British Columbia, June 1989.

Mann, William, & Thompson, Sandra. 1986. "Relational propositions in discourse." *Discourse Processes*, 9 (1): 57–90.

Optimal interpretation for rhetorical relations

Henk Zeevat
ILLC, University of Amsterdam

This paper explores the use of a simple and computational optimality theoretic pragmatics (OT pragmatics, OTP) to the analysis of the rhetorical structure of texts. It introduces this variant of OTP briefly and then shows that it can be applied to explain coherence, rhetorical relations, discourse trees and context dependence using the occasional extra premiss. It thereby improves on existing accounts by reducing rhetorical structure to general pragmatics. It also contributes to the computational problem of inferring rhetorical structure by giving more structure to the inference processes involved, though it is as dependent on better ways of estimating semantic plausibility as any other system.

1. Optimality theoretic pragmatics

Optimality theory (OT) is a natural environment to formalise pragmatics. It gives an account of defaults that is simpler than competing ones and defaults are the bread and and butter of pragmatics. While other formal accounts of pragmatics can likewise exploit the use of defaults in simplifying axioms, the OT framework forces their expression as a small set of very general principles which are applicable throughout pragmatics and thus prevents the development of accounts which only work for isolated phenomena. In this way, whatever one states about e.g. rhetorical structure will have its repercussions on presupposition projection or the derivation of implicatures. In fact, the proposal of this paper on rhetorical structure developed in an organic way by progressive abstraction from the influential account of presupposition projection of Heim (1983) and van der Sandt (1992). It was also heavily influenced by other work on OT pragmatics like Blutner (2001), Beaver (2004), Jäger (2003) and Mattausch (2001a) as well as by the pioneering approach of Hendriks and de Hoop (2001).

Calling the system OT pragmatics merely means that it is conceived as a system of strictly ordered soft constraints. Both the constraints and their ordering should be universal. It will be argued below that it is identical to a constraint system that would give explanations for communicative behaviour of other subjects, itself a special case of the notion of explanation of natural events. This means that the constraint system as such is not part of the development of human

languages and may operate just as well in other species, provided joint attention to common goals and common questions can be assumed. It can even be argued that it is not OT at all, but merely another case where a cognitive problem can be described by a system of strictly ordered soft constraints.

The version of OTP used in this paper is purely interpretational (it selects an optimal interpretation for an utterance) and is thereby comparable to relevance theory Sperber and Wilson (1984). It derives from an unpublished attempt by Blutner and Jäger (1999) to reconstruct the DRT-based presupposition theory of van der Sandt (1992) within optimality theory. *NEW is a generalised version of *ACCOMMODATE and of Hendriks and de Hoop (2001)'s DOAP principle. PLAUSIBLE comes from Mattausch (2001)'s attempt to reconstruct the temporal reasoning in Asher and Lascarides (1993). The replacement of Blutner's STRONG by RELEVANCE is influenced by Van Rooy (2003).[1]

In this way, it is different from the various bidirectional accounts of OT that have been offered (Smolensky (1996) and Blutner (2001) are the original versions). These accounts assume a single constraint system that can be used to select the best forms for an interpretation and the best interpretations for a form by running a match between possible forms and possible interpretations respectively. A normal phenomenon in a system of this kind is that *F* can be the best form for interpretation *I* without *I* being the best interpretation for F. In such situations, it seems absurd to use *F* for *I*: one is guaranteed to be misunderstood. Or inversely, *I* can be the best interpretation for form *F*, but *F* is not the best form for *I*. Here, the interpretation is spoiled by the thought that one would never have said it that way oneself and consequently by not having a proper explanation for why the other speaker said what she said. Strict bidirectionality (strict BIOT) outlaws these situations: if *X* wins for *Y*, *Y* must also win for *X*. Real winners for an input are the ones for which the input is also a winner.

Strict BIOT can be weakened to weak bidirectionality (weak BIOT). The reason for wanting to do so is that strict BIOT rules out any pair $\langle X', Y' \rangle$ where X' is more marked than *X* and Y' is more marked than *Y*. Such a pair is ⟨*cause to die*, **kill in an unusual way**⟩, which is eliminated by the pairs ⟨*cause to die*, **kill in an unusual way**⟩ and ⟨*kill*, **kill in a normal way**⟩, since *cause to die* is more marked than *kill* and **kill in an unusual way** is the more marked meaning.

1. The formalisation of Van Rooy by means of decision problems is interesting and consistent with what I do here. It is however not easy to see how Van Rooy's approach can be made dynamic, i.e. how the influence of the ongoing discourse on the decision problem can be modelled. The current approach is more in line with Rooth (1992) and Zeevat (2006a).

In strict bidirectionality a pair $\langle X, Y \rangle$ is out if there exists an X' or Y' such that X' is better for Y or Y' is better for X. Let's call pairs $\langle X, Y' \rangle$ or $\langle X', Y \rangle$ improvements. A strictly bidirectional pair has no improvements. A weakly bidirectional pair merely has no weakly bidirectional improvements. This sounds like a complex definition,[2] but all strict bidirectional pairs are weakly bidirectional and OT well-orders the pairs, so that a recursion can be set up.

Weak bidirectionality approximates the Horn-Levinson concept of pragmatics and crucially is able to formalise the M-principle, i.e. iconicity.

Unfortunately, there are very serious problems with both notions of bidirectionality. First of all, there is the Rat-Rad problem (in a bidirectional system, the pronunciation /rat/ in German would always be interpreted as the abstract form Rat, while it is ambiguous between Rat and Rad. This problem is not specific for phonology but also crops up in syntax, as shown in Zeevat (2006b) and is a problem for any notion of bidirectionality in which production and generation use the same constraint system, as in the strict and weak bidirectionality introduced above. Beaver and Lee (2003) discuss more problems with bidirectional systems but crucially show that a reasonable system of constraints for Korean syntax, under a weakly bidirectional interpretation predicts a completely unattestable unbounded series of weakly bidirectional equilibria. This argument is lethal for a synchronic version of weak BIOT since it is (again) a direct consequence from the well-ordering on pairs imposed by an OT constraint system.

A third problem is the prediction of fully bidirectional systems that synonymy and ambiguity would occur with the same low frequency: there is considerable and identical evolutionary pressure against both of them. Weak bidirectionality would even seem to rule out both phenomena altogether. Full synonymy however seems a minor phenomenon while the ambiguity of natural languages is overwhelming and is considered by many to be the main obstacle to constructing machines with human-like language capabilities.

The reasoning behind the bidirectional systems is however hard to dismiss. There is considerable evidence that production and interpretation are interconnected and the null hypothesis for explaining the interconnection is that the abstract description of the processes, the grammar, is the same. Second, it would not seem to be the case that one ever interprets an utterance in such a way that the result of the interpretation cannot be seen as a suitable input for generating the utterance oneself (allowing for performance errors and differences in perspective and

2. As shown in Jäger (2000), which also shows that weak BIOT loses the property of simple monodirectional OT with a boundary on the number of errors that if the system is made up from regular constraints, the resulting system is also regular.

competence between the speaker and the hearer). And to use a form for a meaning knowing full well that it will be interpreted in the wrong way seems to border on insanity.

Something has to go. For the pragmatic theory of this paper (see also Zeevat (2009)), the point of departure is precisely the argument for bidirectionality which forbids interpretations that would not explain the use of the utterance by the speaker. This is a corollary of Grice's definition of meaning$_{N\,N}$ in Grice (1957): unless the hearer thinks the putative intention behind the utterance explains why it was made, she cannot think she has recognised the intention behind it. But the relevant notion of explanation is just that the form must be optimal—at least in the speaker's grammar—for the intention, i.e. the interpretation. It seems this principle cannot fail, since it is constitutive of the notion of interpretation as such. Given that for standard constraint systems, selecting the best interpretation does not give the same connections between form and meaning as selecting the form for a given meaning, it follows that a competition using the generating system of constraints cannot be the right model of interpretation.

It is more plausible to think of interpretation as a different process (e.g. one that locally connects the lexically evoked concepts in a plausible way with each other and the context) which is filtered by the generating system. This could be called the mirror neuron model of interpretation: speech behaviour like other observed behaviour by other creatures brings about the same excitations of the mirror neuron system as would occur when the same behaviour would be carried out by the motor system and includes representations of the goals of that behaviour (Gallese (2003)). A model of interpretation of this kind predicts the amazing possibilities of syntactic repair that we seem to have and which allows us to engage in dialogue even when a common language is largely missing or when the channel is quite noisy. And also why language understanding is much more extensive than the language that can be produced. It would also give a model in which interpretation is guaranteed to be in harmony with generation.

Pragmatics proper is about the further filters on interpretation. This paper assumes three more filters: a constraint that maximises plausibility, a constraint that maximises coherence and a constraint that maximises relevance, applying in that order.

Does this mean that bidirection has completely gone overboard? The answer is yes, if bidirection is a formal condition on generation. But it still comes back in three forms, as a condition on interpretation, as a driving force in learning and language evolution and in the form of expressive constraints in OT syntax. The first of these was introduced above: interpretation is dependent on generation by being the strongest contraint on possible interpretations of an utterance. But there is also an important reflex of interpretation in generation. Pragmatics prefers

certain readings and speakers will be misunderstood if they want to express the less plausible, less coherent or less relevant reading in a context. This has led to lexical, intonational and syntactic modes of expression whose primary purpose seems to be to mark against these tendencies. The existence of vast inventories of nouns, verbs and adjectives in natural languages should be attributed to the tendency to go for the stereotypical and the contextually expected (the tendency enforced by plausibility). Adversative and mirative markers seem to serve the same purpose of allowing implausible interpretations: they do not force implausible interpretations but prevent repair of implausible interpretations, since the marker needs to be justified by a less than normally plausible interpretation. Marking new information by intonation, articles and additive particles serves as a counteragent against overly coherent interpretation. Twiddly intonation, particles like "well" and modals mark against unwanted relevance. In fact, most properties of natural languages can be seen as an answer to unwanted interpretations due to pragmatics. And pragmatics itself is nothing more than the theory of explanation applied to communicative events, where a purpose can be assumed in the sender.

Pragmatics helps in selecting the right interpretation, but as such is also the mechanism behind misunderstandings. Linguistic expression used for a meaning *A* but understood as *B* will not replicate well in language history and other expressions for *A* will invade. Pragmatics accounts for why an expression for *A* will be interpreted as *B* and thereby is the account of one of the two driving forces behind language change: the need for functioning expressive power. The other is phonetic, phonological, morphological and syntactic erosion which finds its explanation in the difficulties of upholding norms of speaking in societies of speakers, in the absence of clear and conscious criteria.

The same driving force of expressive power can be seen in synchronic syntax in expressive constraints forcing the expression of certain features of the input. Why is an occurrence of "too" normally obligatory? Because otherwise, the utterance would be interpreted too coherently, by identifying the earlier and current state of affairs. Why does the reference to John come out as "he"? Because otherwise the interpreter may think that it is not the same person as in last sentence and may misunderstand the name. These expression constraints are bidirectional: they assign errors to a candidate iff it is also optimal for the input where the feature is toggled. They are constraints that are freely ranked with the others, though the order is not completely arbitrary since the features have an inherent communicational importance. They can have grammaticised to a particular morphological, lexical or syntactic device for their expression (e.g. tense, plural, imperfective), but can also employ a number of expressive devices (e.g. subject in Dutch or Russian which can be marked by case, agreement or word order). Expressive constraints are visible expressions of bidirection in syntax. They make the speaker's

task easier of guaranteeing that she will be properly understood. But the fact that such constraints did emerge and grammaticise makes it questionable that speakers are really able to guarantee that they will be understood. The problem to which expressive constraints seem to contribute would not have existed if speakers were able to avoid misunderstanding altogether. At the same time, the emergence of expressive constraints also shows that speakers try to avoid unwanted interpretations. It is clearly a hard task. The importance of feedback mechanisms in natural dialogue seems to be evidence that speakers and hearers can and do not count on perfect understanding.

2. Pragmatic constraints

Faithfulness in optimality theory refers to the relation between the input and the output. The concept does not come out in the clearest possible way in phonology where it is customary to use the Latin alphabet both for the abstract input phonemes (defined in the lexicon) and the more concrete phonemes that form the basis for pronunciation: faithfulness then seems to be identity, though it really is not. In syntax, there can be no identity. While useful versions of OT syntax can be given where the input is rather linguistic, for pragmatics, it is necessary to take actual speaker intentions as input. These intentions have internal structure and various complex or primitive constituents of the intentions can have features, in virtue of what they are or in virtue of a relation they bear to the context. If a feature of the input or a feature of a constituent is expressed by a candidate realisation, the candidate is faithful with respect to the feature, otherwise it is not. If one could list the relevant features and characterise speaker intentions by sets of them, this would be an easy issue.

But there are good reasons for doubting that there is a universal inventory of features that get expressed in language and even more reason to doubt that intentions can be characterised as sets of such features. To start with the last problem, intentions should contain the content of what the user wants to express and this can be arbitrarily complex. It follows that intentions can only be finite if a limitation is adopted, e.g. to intentions corresponding to simple clauses. It may then be possible to pack the lexicon, semantic and contextual properties into a single set of features. But a limitation of this kind is not plausible if intention recognition should also allow the interpreter to reconstruct the reason of the speaker for producing the speech act in the first place, something which seems unavoidable in pragmatics.

The second problem is that the content of lexica of different languages can be full of idiosyncrasies. Assuming a universal language of thought does not really

help here, since the lexicon of the language may well be unable to express aspects of the universal thoughts. And it is not clear at all that linguistically important semantic features (e.g. the feature that should control negative polarity items, or animacy) allow of a universal definition.

The view of this paper is that the input of the speaker is the speaker's intention. It is given by the speaker's goal in the conversation, together with what the speaker knows about the language, the worlds and the context. The intention as a representable entity comes into being as the sum of the decisions that make up the production process. These are forced by the constraint system governing the production of natural language utterances and may partially depend on properties of the inventory of the language. It follows that understanding is not so much that the hearer obtains the same intention, but that the hearer forms a picture of the intention that would allow her to make all the decisions in the same way, if she would have been the speaker and allowing for differences in competence and error.

On this view, correctness of understanding is dependent on the language used as given by the inventory and the constraint system. Intention recognition can be limited to understanding those aspects of the goal of the utterance and the information that the speaker had at her disposal that played a role in shaping the utterance.

FAITH is the strongest constraint in the constraint system of this paper. It tells the hearer to select those interpretations that for the speaker could constitute a reason for making precisely that utterance. It involves the reason why the speaker is speaking and everything connected with making decisions that determine the form of the utterance. The speaker may repair the utterance in view of performance errors or incomplete competence of the speaker. FAITH however minimises the number of such repairs. A selected interpretation should not have competitors which would require less error correction.

Interpretive accounts of natural language semantics like Montague Grammar, Discourse Representation Theory, Head Driven Phrase Structure Grammar or other grammar formalisms that integrate a treatment of semantics are approximations to FAITH. The same holds for versions of OT syntax that take as input a semantic representation of some kind.

They are approximations only, because they do not attribute to the speaker a reason for speaking and because they do not contain an account of non-literal use of language, such as irony or metaphor. An OT syntax which would start from intentions as input and which would incorporate the use of non-literal language use would be a better approximation, but treatments have not been formulated yet.

For natural events, actions by humans and other organisms and non-linguistic communicative acts, the corresponding demand on interpretations would be that the interpretation explains them.

PLAUSIBLE is concerned with consistency and likelyhood of the interpretation. It rules out an interpretation that has a more consistent or more likely interpretation as a competitor. It can be related to the use of consistency and likelyhood in various linguistic processes, like e.g. in the pragmatics proposed by Gazdar (1979) or in ambiguity resolution as practiced in many current lines of work in natural language processing.

Outside language, it compares explanations for consistency and likelyhood. An explanation is more likely to the degree the cause it proposes is itself likely and to the degree the cause is known to bring about the effect that has to be explained.

*NEW minimises the number of new objects by the interpretation, but also militates against new unanchored objects and changes of syntactic role of the object. Plausible proper interpretations that have more new objects or less anchored objects or that retain fewer objects in their syntactic role are rejected in favour of equally plausible and proper interpretations which do better.

Outside language, in scientific explanation. the principle is just Ockham's razor. It is not significantly different in everyday explanation. The simpler the explanation, the better.

RELEVANCE is closest to the Gricean maxim. The interpreter has the right to expect that the speaker addresses issues and questions of which it is common ground between her and the speaker that they are of interest to the subjects of the conversations. An issue or question can start out as such, but dialogue can add new questions and answer others. Most overtly this can be done by asking questions and stating goals and so raising the question how the goal can achieved. But another important other mechanism is by activating propositions that are not yet decided, stating surprising facts, mentioning new objects etc. In all these cases, questions are added: is the activated proposition true, how did the surprising fact come about and how did the speaker know about it, who or what is the new object? RELEVANCE maximises the number of activated questions that are answered by the utterance, by matching the utterance with the activated question whenever that is possible and then adding the assumption that the information provided in the utterance is all there is to know about the question.

There is no corresponding principle in the explanation of natural events since there is no justification for the idea that natural events occur in answer to a goal the interpreter shares with nature. But there is the same principle for non-linguistic communicative behaviour. If you show me a photograph in a common ground where we want to know what happened at the departmental party, I will take you to imply that the picture is about that party. If we are skating over the lake and you shout "aargh", I will draw the conclusion that you want to warn me for a natural peril in the situation.

The ordering has to be exactly this way. Without minimisation of new objects, it would be easy to have too much relevance. Without plausibility to constrain the interpreter, she would identify everything. Without FAITH in overall control, plausibility would limit the hearer to trivial messages only.

3. What should a theory of rhetorical structure achieve?

In the first place, it should be able to support models of text generation. The text planning and accumulation modules distinguished in Reiter and Dale (2000) should be guided by the ideas of what is possible structure according to the account of rhetorical relations. There is a strong case for developing suprasentential OT syntax but this has not been done yet.

Second, it should similarly support models of comprehension. An important role of rhetorical structure theory in combination with theories of information structure is to give criteria for comprehension. Information theory should give an account of the purpose of a move in a dialogue. Current theories seem to contribute the requirement of a maximally strong discourse relation between the current DCU[3] and the pivot DCU and constraints on the identification of the pivot.

The literature on rhetorical structure contains several proposals: the abduction model of Hobbs et al. (1990), the glue logic of Asher and Lascarides (2003), the "greedy parser" of Prüst et al. (1994) and others. The theory given in this paper has one contribution to make here. It distinguishes three kinds of defaults (those coming from PLAUSIBLE, *NEW and RELEVANCE) and adjudicates on the resolution of conflicts between these 3 kinds of defaults. In this way, it is more structured than Hobbs or Lascarides and Asher where all the defaults are in direct competition and where weighting or specificity is the single adjudicator.[4]

Second, rhetorical structure theories have powerful predictions to make on the interpretation of anaphoric elements such as pronouns, ellipsed constituents, proper names, descriptions, nouns, tense, particles, (implicit) (temporal) locatives, factives, (pseudo-)clefts, predicates with sortal restrictions, intonationally marked

3. DCU (discourse constituency unit) is terminology introduced by Polanyi (1985). It refers to the constituents that enter into rhetorical relations and covers subsentential as well suprasentential units.

4. It may be that there is mileage in the proposals of Zeevat (2008) which proposes a cascade of procedures (one for each constraint) that try to further instantiate an underspecified representation. In that case, default processing can be limited to a mechanism like default unification. But it is really too early to tell: a crucial ingredient is a formal treatment of non-literal language use and better ways of estimating plausibility.

topics, contrastive stress, additive marking etc. These predictions come from the identification of the pivot and the relation the DCU bears to it.

The strongest claims are made by Prüst et al. (1994) where pronominal resolution and VP anaphora are reduced to a mechanism that bears a certain resemblance to mechanisms that have been proposed for the computation of "discourse topic". The mechanism enforces a maximal parallelism between two DCUs. The most specific common denominator of a semantic representation (or perhaps more properly a hybrid representation between syntax and semantics) of the pivot and the representation of the DCU is computed. *mscd*(*A,B*) is the most specific generalisation of *A* that still unifies with *B*. If *A* is instantiated where *B* has variables (i.e. pronouns or pro-VPs) *mscd*(*A, B*) has the values of *A*. If *A* and *B* have conflicting instantiations, the corresponding place in *mscd*(*A, B*) will have a variable which will be instantiated by the value that *B* has there. The unification *mscd*(*A, B*) $\wedge$ *B* gives the result of the interpretation process for *B*.

The mechanism is important, but limited to a subset of the phenomena. While it can be extended to N- and VP-ellipsis, some particles, implicit locatives, tense and contrastive stress, it does not seem to be applicable to proper names, nouns, descriptions, additive marking, other particles, factives and sortal restrictions, because these cannot be described as paralelism effects and do not necessarily have an antecedent in the pivot. The mechanism also does not give the right result in discourse relations that do not require parallelism (see Kehler(2002) for extensive discussion): parallelism is only required in the case of Contrastive Pairs, Lists and Question Answer Pairs and optional with other relations. But—also in the absence of parallelism—rhetorical structure has an important contribution to make to cross-sentential resolution by identifying the pivot which binds pronouns, tense and ellipsis in the new DCU.

4. Pragmatic constraints in rhetorical structure

It is particularly easy to see that PLAUSIBLE has an important role in rhetorical structure. In fact, this can be seen as the dominant view among the theorists: enough competing default rules of various strength or specificity should do the job as in Hobbs et al. (1990) or Asher and Lascarides (2003).

In (1) a pivot sentence *John fell* is combined with three new DCUs: *Mary pushed him, Mary smiled at him and Mary hit him*. All three DCUs can be marked for their connection to the pivot by explicit markers, here: *but* (Concession or Formal Contrast), *then* (Narrative), *so* (Result) and *because* (Cause). Typically, the absence of a marker leads to the most plausible connection given the predicates. In (a) Cause, in (f) Narration and (k) nothing is chosen because it could be

anything: Cause, Narration, Concession or Result. It seems clear in (k) that normally a choice will be based on the particular context where e.g. John is doing a balancing act so that being hit is a plausible cause of falling, John and Mary are in a fight where it is plausible that Mary seizes her chance of hitting John when he falls, or sporting expectations are set up that would normally prevent Mary from hitting John when he falls, or the reverse, where Mary has threatened John to hit him if he would be so clumsy as to fall.

The markers override any of the defaults and may effect reinterpretations of the context to make it possible to have concessive or causal interpretations. This is only possible with explicit markers that are interpreted by FAITH. If the context cannot be reinterpreted to allow the concessive, causal or result interpretation, pragmatic incorrectness results. The markers seem clumsy if they do nothing more than confirm the most plausible relation anyway. They should be left out in that case. At the same time, their occurrence is not optional if the speaker intends a particular relation while PLAUSIBLE would give a different relation (though a different means of expression may be chosen). The force behind the need to mark can be expressed by an expressive constraint (that would also be satisfied if the connection is clear by plausibility).

(1) a. John fell. Mary pushed him.
b. ?John fell. Because Mary pushed him.
c. John fell. Then Mary pushed him.
d. John fell. So Mary pushed him.
e. John fell. But Mary pushed him.
f. John fell. Mary smiled at him.
g. John fell. Because Mary smiled at him.
h. ?John fell. Then Mary smiled at him.
i. John fell. So Mary smiled at him.
j. John fell. But Mary smiled at him.
k. ?John fell. Mary hit him.
l. John fell. Because Mary hit him.
m. John fell. Then Mary hit him.
n. John fell. So Mary hit him.
o. John fell. But Mary hit him.

*NEW applies to all discourse markers in the new DCU. This is a rather large class if one takes the criterion for being a discourse referent that of being a possible antecedent for some kind of anaphora or ellipsis. This is not necessary. One could let the criterion be the existence of overt anaphoric elements, but that would not be a universal criterion since elliptical anaphora is a typological option. On the other hand, for relations or topics (arbitrary abstracts) overt and specific devices seem to be rare.

This gives a list like (2) for *NEW to apply to.

(2) objects
kinds
moments of time
sets
events
states
facts
thoughts
spatio-temporal regions
relations or topics

Relations or topics require some argument. Consider (3). (3a.) takes an antecedent *X gave Mary flowers*, (3b.) *X gave Y flowers*, (3c.) *X gave Y Z* and (3d.) *John gave X flowers*.

(3) a. John gave Mary flowers. Bill did too.
b. John gave Mary flowers. Bill Sue.
c. John gave Mary flowers. Bill Sue chocolates.
d. John gave Mary flowers. And Sue.

It is not necessary to think of elements of the last category as being created by some construction algorithm. It is enough that they are available for binding ellipses and that they can have levels of activation. But both of these properties can be derived from the antecedent utterance itself: it may have a level of activation which makes an abstract contained in it suitable as the antecedent of a certain ellipsis or overt element and the abstract itself can be derived from the utterance when needed. The description of the abstract as the topic should be underpinned on intonational grounds. In fact, Rooth (1992) notices that contrastive intonation on Mary and Sue or on John and Sue leads to quite different interpretations. In the corresponding *John gave SUE flowers too, John gave X flowers* is destressed and thereby marked as given, something the interpreter needs to check.

Let's try to reformulate *NEW appropriately. It should generally always prefer old over new, highly activated over lower activated, parts over merely related. If in addition preservation of certain linguistic features (e.g. AGENT, THEME) is preferred over changing them from the antecedent, maximal parallelism becomes the interpretational norm. The additional demand makes sense under the interpretation of *NEW as a perceptual principle of conservatism: when there is no new information assume everything stays the same. Adding this principle recreates the MSCD-based mechanism of Prüst et al. (1994) while avoiding the limitations because *NEW is a soft constraint.

RELEVANCE would be mainly responsible in rhetorical structure for the strengthening of discourse relations. If e_2 is contingent on e_1 as in Narration, e_1

addresses the question *what caused* e_2?. RELEVANCE instructs the interpreter (if e_1 is a plausible cause of e_2) to take e_1 as the answer which changes the discourse relation into Result. Other strengthenings of this kind are Background to Cause, Background to Justification, Reformulation to Conclusion.

This leaves FAITH. FAITH is first of all responsible for marking devices which can suspend the workings of PLAUSIBLE, *NEW and RELEVANCE. But if its proper formulation involves reconstructing the speaker intention as is assumed in this paper and not just the projection of lexically and syntactically expressed meanings into a semantic representation, it is responsible also for the motivational aspect of the utterance: the hearer must find a reconstruction of the intention that makes it clear not just why the speaker decided on these words and constructions but also why it is worthwhile to make the utterance in the first place. And—where the utterance is complex—also why the speaker thinks the subutterances are worth making.

5. Coherence

While it is necessary to allow for errors and lapses, it is rational to assume that a speaker has a reason for speaking and that there is a reason behind any part of her utterance. Some of the parts find their reason in an overarching reason for producing a larger part: the reason is just that they contribute to the larger enterprise. Thus the occurrence of a word like "he" in a larger part of speech like "he ran away" can be explained by whatever the reason is for the larger utterance: without the "he" the speaker would not identify the event to the necessary degree which presumably would defeat the intention behind mentioning the fact that "he ran away".

But this changes in those cases where the degree of freedom is greater: for optional modifiers, participials, and for extra separate sentences. Now it could well be countered that for such extra material the hearer can find an independent intention behind its utterance. But this does not work: the hearer faces the double task in this case of explaining not just the extra material but also of explaining why it is syntactically integrated with the other sentence or appears adjacent to it -in the absence of a sign that makes it clear that the speaker intends to shift or interrupt the current course of the conversation, as in (4).

(4) John is away. Now for something entirely different. Somebody took my cup. Do you know anything about it?
A: What time is it?
B: 4.30 and why did you not show up yesterday?

The fact of the matter is that the best explanations of the extra material connect the extra material by a discourse relation to the other sentence. It is only if the speaker rules out such a connection or if it is not possible to fit in the new material with the drift of the conversation until the current point, that shift can be assumed.

The proper explanation of the connectedness of a new utterance or an optional part of an utterance starts by denying that it needs to be. Not every utterance/optional material is connected to a pivot (it is not necessary here to assume pseudo-relations like interruption or topic change). Then FAITH demands an explanation for optional material and in particular for extra sentences. *NEW finally brings about a preference for given topics, given objects, times etc. This connects the explanation for the unaccounted material with the current drift of the conversation.

A quite tempting direction that has been taken here but still has not been exploited to the full extent is the perspective of generation systems. In systemic grammar (the framework to which the first work on rhetorical relations (Mann and Thompson (1985)) belongs) the generation process is conceptualised as a series of connected choices the speaker has to make and provides decision procedures in some cases (e.g. obligatory marking of tense, choice of article for NPs). This gives a large range of explanations for the choices of a speaker. But the pragmatically most interesting choice points are the ones where choices are not forced. FAITH seems to force finding a reason for any choice that is not arbitrary. As has been noted by Dale and Reiter (1996) for some particular cases, some of these choice points are connected with a range of pragmatical implicatures. In this view, unexpressed rhetorical relations are implicatures of the choice points that involve the insertion of optional material and of choices at the text level.

(5) The angry farmers blocked the road.
The farmers -angry because of falling prices for their products- blocked the road.
The farmers blocked the road. They were angry.
possible implicature: the farmers blocked the road because they were angry.

The implicature may be the joint effect of FAITH, *NEW and RELEVANCE, where FAITH demands that there be a reason for the extra material, *NEW that it is strongly connected to the pivot, and RELEVANCE determines whether the cause of the blocking is at issue (it might also be Background).

6. Rhetorical relations

In the following, it is attempted to answer the question why there are the rhetorical relations that researchers like Mann and Thompson (1985), Hobbs (1979),

Grosz and Sidner (1986), Polanyi (1985) and others have found. It does not do to explain these as cultural artefacts that have proved useful in conducting conversation and writing texts. This would predict substantial variation among cultures in the inventory of relations and nothing really spectacular has been found in this respect.[5]

More promising is to see DCUs bearing a rhetorical relation to a pivot as specialised speech acts that follow up on other speech acts. Turn changing speech acts like answering, accepting, rejecting have a relation to the other turn and cannot be understood without knowing the content of the other turn. In rhetorical relations, the turn does not change and the relation is to an earlier element of the same turn, the pivot, without which it cannot be fully understood.

While it is right to see the DCU as a special kind of assertion and the relation it has to the pivot as part of what makes it a special kind of assertion, it does not seem this perspective throws light on the question where the inventory of relations comes from. The major speech acts themselves seem to arise naturally out of the sort of things one can do with language: give information, ask questions, enter into commitments etc. Specialised speech acts seem to come from the defaults associated with their superclass and lead to marking devices for indicating the special case. E.g. a default assertion expresses knowledge of the speaker and addresses an unsettled issue in the strong sense that there is no bias for or against any particular way of settling it. This guarantees that the interlocutor can just accept the content of the assertion, unless she has conflicting information. Special cases need to be marked (as illustrated in (6)): if the speaker has merely inferred it, if it is only a suggestion the speaker is making, if negative or positive evidence is present or if the speaker needs confirmation from the interlocutor.

(6) John has left.
John must have left.
John may have left.
John has indeed left.
John has left after all.
John has left, hasn't he?

The situation with rhetorical relations is not different. There is a default with respect to what is the pivot (the last simple DCU) and a default with respect to

5. There are different strategies in telling stories, explaining complicated states of affairs, different politeness norms, but not clearly different rhetorical relations. If the treatment of these strategies can be achieved in an extended OT syntax, this would predict the existence of a typology of rhetorical strategies. If the relations are explained —as in this paper— from the theory of interpretation, this predicts the same inventory of relations.

what the DCU does (the same thing as the pivot). All other things are marked and need to be protected from misunderstanding by connectors, particles, lexical material and intonation.

This section tries to argue that rhetorical relations can be explained from *NEW. By default, the pivot and the current DCU are completely the same. Rhetorical relations classify the transgression of *NEW. Strengthened relations are obtained by RELEVANCE.

To make the point, it is helpful to assume that DCUs can be represented by four parameters: a list of participants, a spatio-temporal location, a relation and a segment topic (the last parameter is not defined in an opening DCU).

The segment topic is part of information structure, but is distinct from the DCU's own topic. When the DCU's topic can be seen as part of the segment topic, one obtains the tripartite view of information structure proposed by Vallduvi (1992). The segment topic is a proposition (it can be seen an issue to be decided) for the settling of which the current DCU is relevant.

Given *NEW, the most unmarked next DCU is a full Repeat: everything is preserved, participants, relation, location and segment topic. But a repeat is only useful, if it is likely that the hearer has somehow failed to recognise the utterance well enough.[6]

A full Repeat would normally lead to a conflict with FAITH: if the DCU is made and accepted, any of the goals for which it was produced in the first place should now have been reached and, if they are not, it cannot be expected that merely repeating will achieve them. So the speaker cannot reach any goal by a mere repeat and the hearer is consequently not able to reconstruct an intention behind it.

For a non-repeat, the interpretation best meeting *NEW is one where the parameters shift as little as possible. Here the strongest similarity is given by retaining the main participant(s)[7] and the location. This forces a closely related sentence topic and a compatible predication. The rhetorical relation is known as Reformulation. Jasinskaja (2007) shows that with default intonation and without

6. Or for getting attention back to an earlier part of the exchange: Reminders, another purpose of later repeats can be to remind the interlocutor of an earlier commitent, which she now seems to give up (Reconfirmation). Yet another purpose can be to use old information to explain or motivate the pivot (Justification). These Repeats are often marked for their new function and in all these cases the pivot is not the repeated DCU which would occur much earlier.

7. A precise definition is difficult. The examples in (7) are covered by making it a demand that one discourse referent of the pivot is maintained in addition to the location.

explicit markers, Reformulation is the most preferred interpretation: it redescribes what happened to the participants at that time and place.

A new sentence without an explicit marker is however not guaranteed to meet FAITH and PLAUSIBLE. FAITH can change participants (by full NPs) or change location by temporal modifier and locatives, PLAUSIBLE can force change of location and participants if the identities are hard to swallow. The second clause of (8) redescribes what befell Alena (from Jasinskaja (2007)).

(8) Alena broke her skis. She lost her only means of transport.

Time, place and participants are the same. The only change is the predication and the point is the entailment: Alena no longer has a means of transport and stuck to her location. In (9), there are problems with assuming that Alena and the location are maintained: is Alena John? Are the skis a car? The implausibility forces the interpreter out of the Reformulation assumption.

(9) Alena broke her skis. John smashed his car.

Change of location (change of time or place) can be divided into proper change to a disconnected spatio-temporal region and change to a subregion or overlapping region. Change to a subregion is typical of elaborations. Overlapping regions are typical of causal connections. Disconnected regions lead to distinct sentence topics.

A similar division can be made with respect to participants. A participant can be continued or its role can be occupied by a subparticipant (a single individual or proper subset of the group that was the original participant, a part of the participant or a subquantity of a quantity participant) or can change to a distinct object. Change to a subparticipant indicates an elaboration. Change to a distinct participant a distinct sentence topic.

Changes to subregions and subparticipants indicate elaborations and this is the default if Reformulation cannot be assumed. In this case, the pivot becomes the segment topic. It now functions as the topic for the whole elaboration. The typical elaboration strategy for dealing with a topic is to break it up in distinct parts by breaking up its location or one of its complex participants and treating the parts in turn.

Making the pivot the segment topic does not mean that the current DCU has the pivot as its sentence topic. It merely means that the pivot is not abandoned and is still necessary for further semantic processing, e.g. for exhaustification effects. The same is also going on in discourse relations like Explanation, Background, and Justification. But this needs some explanation.

The explanation in fact takes care of some other questions as well. If our model of addressing topics is question answering, i.e. addressing the topic is

giving an answer that settles the topic, it would not be understandable why there are relations like Elaboration and Restatement in which an already settled question is readdressed. What could be the point of that? But the model of question-answering seems to be misconceived. An assertion is not offering a proposition for belief, it offers a proposition that appears to the speaker as knowledge, something she has grounds for accepting as reliable. The purpose of communication of facts is not to let it be known what the speaker knows, but to construct the knowledge in the interlocutor so that she knows these facts as is necessary for her purposes. That is why the details given in an Elaboration and in a Background matter: they tell the interlocutor how the knowledge presented itself. In a causal Explanation, the speaker also underpins the proposition, the explanation makes the truth of the proposition understandable. In a Justification, the speaker states the grounds for accepting the proposition as true, grounds which may be sufficient for the interlocutor as well. In elaborating Lists and Reformulations, the segment topic helps to determine the sentence topic and consequently the exhaustivity effects. In Explanation, Background and Justification, the segment topic gives the issue that the DCU helps to settle. The segment topic is necessary in order to state the causal and inferential connections.

So far, the pivot is kept. Quite literally in a Repeat or by keeping it as a segment topic as in Reformulation, Elaboration, Background, Explanation and Justification.

In the other cases, the pivot is discarded, a violation of *NEW. In these cases, the location or the main participant(s) is distinct from the location or participant(s) of the pivot. If there is a segment topic, this is maintained in List and in Narration relations. In a List relation, the sentence topics are distinct subtopics of the segment topic, where the subtopics are typically given by splitting up a participant or the spatial location. In Narration, the division is by temporal location. There are however two properties of a Narration that makes Narration different: the fact that time moves forward in a Narration and the fact that successive event have to be contingent on each other. This mirrors the stream of experience (or the structure of plans) and cannot be reduced to abstract pragmatics. Event descriptions themselves seems to move the attention from the start of the event to the point where it happens[8] Again this seems iconic with the experience of an event. From these considerations, it is perhaps not surprising that Narrations can have formal

8. Movement of time is intimately connected with the aspectual systems one finds in many languages. Perfective marking can be described as locating the event at a given point to reach another later point in time. If another perfectively marked event is located there, the contingency relation that for Hobbs (1979) is definitional of Narration is nearby, since the resulting state of the first event should hold when the second event occurs, thus making it part of the circumstances that allow the second event to occur.

properties that set them apart from other relations, like specialised tense forms (past tense on non-stative verbs in Dutch) or zero subjects as in Chinese (which are limited to subordinate sentences otherwise). Nonetheless, Narrative sequences are still a special kind of List given by division of a segment topic by splitting up the temporal location.

The most marked cases from the perspective of *NEW are Contrastive Pairs and Concession. In both cases, the sentence topic of the first DCU (the concession) is a subtopic of the sentence topic and affirms it. The second DCU denies the remainder of the topic, often by entailing or implicating the falsity of the remainder.

Concession is the special case. The truth of the concession in a Concession is a reason for thinking that the whole segment topic is true. The second clause corrects this. In a Contrastive Pair, the causal connection between the first clause and the falsity of the second is not given. For contrast, it is sufficient that the segment topic contains both a part that is affirmed and a part that is denied, as is argued in Umbach (2001).

Contrastive Pair is always marked by two intonationally prominent constituents in each clause, if not by a contrastive marker. It is even harder to make Concessions without an overt marker of concession.

While it is possible to have explicit antecedents for the segment topic of Contrastive Pairs and Concessions (normally a question), this is rare. It is also not necessary, since the pairs give enough information to reconstruct them. In a formal version of this material, one should decide between representing such Contrastive Pairs and Concessions by inserting a zero pivot, that creates the segment topic or by assigning an extra superordinate segment topic to the pair.

This section tried to show that *NEW structures the inventory of discourse relations. While the pivot itself may deal with a wider topic, the default is to go on with communicating the experience reported in the pivot. Within this default, there is a further default of preserving the location and main participant, followed by parts of that and followed in turn by Cause and Justification. If the pivot is abandoned, the default is to retain the segment topic, i.e. keep on doing what the pivot was doing. This default is broken in contrasts and concessions.

So-called discourse popping is not a rhetorical relation. It is a breach of the default that the pivot is the last simple DCU. Discourse popping results if the last simple DCU cannot be the pivot, i.e. if the utterance cannot be constructed to be on the issue of the last simple DCU or contributing to the segment topic of the last simple DCU. In that case, the next candidate for the pivot is the segment topic.

The contingency that is typical of Narration can be strengthened into causality (Result). In the set-up of this paper this is a question of addressing the question what caused the second event, a question that would be activated by any event report. RELEVANCE then lets the interpreter assume that the first event is the

cause of the second. All that is needed is that the first event is a plausible cause of the second. This reconstructs Result. It is natural to assume that pragmatic strengthening by RELEVANCE is responsible for other discourse relations. Narration is a strengthened form of List (is the DCU contingent on the pivot?), Explanation or even Justifiation are strengthened forms of Background, Conclusion is a strengthened form of Reformulation (taking a List as its pivot), Concession a strengthened version of Contrastive Pair.

This section tried to argue that *NEW is the only default in inferring discourse relations and that the basic classification of the rhetorical relation a DCU bears to its pivot is a classification of the ways *NEW is transgressed and obeyed. The full classification of rhetorical relations involves RELEVANCE which is responsible for inferring the strongest relationship, if this is allowed by PLAUSIBILITY. At the same time, overt marking, obligatorily resolved elements (all part of FAITH), contingency estimates, estimates about possible causes and reasons, and estimates about what can addresses which topic (PLAUSIBILITY) have an important role to play in the actual processing. While most of these allow of a computational treatment, PLAUSIBLE is an exception as long as good empirical approaches to plausibility estimation are not available. These seem to be within reach however.

7. Discourse trees

The hierarchical structure normally assumed in accounts of rhetorical structure comes out in the current approach. A DCU dominates all DCUs of which it is the segment topic. A complex DCU is a maximal sequence whose members share a segment topic or a Contrastive Pair. Interruptions are not in the tree, though they may have a tree structure themselves. The tree structure can also be broken by full shifts of topic. This seems an improvement on various approaches in which everything needs to be integrated in a tree, even if they are incoherent by definition. The approach is also not committed to see the tree as the object computed in discourse processing: it just computes the relation between a new DCU, finds the pivot and computes the integration of the new DCU with its pivot. The outcome determines the tree, but the tree is not itself important for interpretation or processing.

The right frontier constraint is a constraint on what the pivot can be. By default, it is the last simple DCU. If that does not work, the more complex DCU terminating at the new DCU can be considered to be pivot. Moving to the segment topic of an unsuitable candidate is the next step and these moves can be iterated. The procedure thus follows the activation patterns. Segment topics of segment topics

are less activated. The last element processed is more activated than the sequence or pair to which it belongs.

8. Context dependency

The theory of Montague (1974) and Kaplan (1989) makes model-theoretic interpretation of an utterance dependent on a set of context parameters. The ambition of current dynamic semantics is to improve on that by including an account of how the utterance changes those parameters. Concentration has been here on one parameter in particular, the information that is available, it being taken for granted that somehow the value of the other parameters can be recovered from that. There are ways of doing that.

It is not sufficiently realised that accounts of rhetorical structure have an important role to play here. They make the interpretation of the current utterance dependent on the pivot as the major source of the context parameters and constrain the choice of the pivot. Moreover, the dependency varies with the relation of the utterance to the pivot, and it is the task of an account of rhetorical structure to explain how it works in particular cases.

The system described in this paper has no pretence of being a mechanistic account that fully determines how the contextual parameters influence the interpretation of the utterance. From an abstract perspective, the opposite appears to be true. It is quite easy to construct situations in which formally correct language use does not allow the context to fully determine the interpretation of the utterance.

But one can hope that in actual language use this only happens by mistake and that full determination of content is not a question of luck but the aim of language users, using syntax as well as PLAUSIBILITY, *NEW and RELEVANCE. Language users can aim for that situation because they are interpreters themselves and so can estimate the degree to which the process will result in the recognition of their intention. In fact, they would use a bidirectional filter. It would be far-fetched to call the proposals of this paper a logic of pragmatics, but it is a way to extend the Montague-Kaplan proposals to the full range of context dependency and to incorporate the task of determining the context for the next utterance.

9. Conclusion

Apart from providing another application area for OTP(Zeevat (2007) (presupposition), Zeevat (2009) (implicatures and pronouns)) and thereby giving an account of rhetorical structure in general pragmatics rather than treating it as

an area of its own, this paper makes a number of points particular to rhetorical structure. This is the list.

1. Both Hobbs et al. (1990) and Asher and Lascarides (1993) and Asher and Lascarides (2003) are approaches to rhetorical structure in which a general theory of common sense reasoning is extended with axioms for rhetorical structure. The approach in this paper largely vindicates that strategy by giving a central place to plausibility. The nature of the rhetorical structure processing is however conceived in a different way by strictly ordering the application of the four constraints. This makes PLAUSIBLE, *NEW and RELEVANCE produce defaults and makes it impossible that the defaults produced by the higher constraints can be overridden by the lower ones. This is computationally simpler.

   The approaches of Hobbs and Asher & Lascarides[9] can also be described as "plausibility reigns supreme". Or better, since these authors assume prior semantic processing, "plausibility reigns supreme after FAITH". In this paper, it was shown that a whole range of rhetorical structure defaults follow from *NEW and RELEVANCE and that plausibility is merely a filter. This seems an improvement.
2. An important innovation is concerned with pronouns and other items that are obligatorily resolved. Pronoun resolution is rightly regarded as pragmatic core business: it cannot be done in syntax or in the semantic composition rules and it is governed by defaults and heuristics. In the OTP of this paper, the actual resolution is separated from the necessity of resolution. It is easy to think that *NEW has something to do with pronoun resolution (its precursor DOAP stands for "do not miss anaphoric possibilities"), but it would be quite unable to make the resolution of pronouns obligatory. In fact, it would be better not to resolve if the resolution would result in implausibility, as e.g. in corrections that are maximally implausible in the contexts that warrant them.[10] This means that the need to resolve a pronoun is due to a syntactic rule that can and therefore must realise a highly activated discourse referent with a personal pronoun. The recognition of a pronoun is therefore incomplete without assigning it a highly activated discourse referent as its referent. The fact that pronouns need to be resolved is therefore part of FAITH. If there is more than one highly activated discourse referent, the decision between them is constrained by PLAUSIBLE. The need to resolve pronouns and other anaphoric items is a powerful factor in recognising rhetorical structure.

9. Asher and Lascarides (2003) reject their earlier assumption of default discourse relations.

10. I owe this point to David Beaver p.c. commenting on a draft of Zeevat (2001).

3. Rhetorical relations are not given by heaven but classify different transgressions of *NEW and can be strengthened by RELEVANCE. The exceptions are debatably Narration and—much less debatably—Concession. This points to a universal grammaticalisation process for possibly Narration and for Concession.
4. The implementation strategy for OTP is more straightforward than for either abduction or SDRT. One needs a model for FAITH which assigns semantic representations to utterances. Any model is in principle fine, in particular all the existing models in natural language semantics, such as Montague grammar, DRT, or GB inspired approaches. Then one requires a model for estimating plausibility. Good empirical methods are still missing, but it is reasonable to expect progress here. Counting discourse referents and investigating their status is trivial. The current proposal for relevance is not difficult to implement if one formulates rules for activating questions. Part of these rules are given in the literature on natural language generation and others can be added (e.g. the questions that presupposition triggers activate and that cause accommodation in Zeevat (2007). Others may be connected to plausibility: low plausibility information naturally raises questions of cause and justification. The algorithm would eliminate successively candidates supplied by FAITH by means of plausibility, *NEW and RELEVANCE.

Alternatively, the algorithm could operate from an underspecified representation coming out of the model for FAITH, to which PLAUSIBLE, *NEW and RELEVANCE try to add extra information.

References

Asher, N. & Lascarides, A. (1993). Temporal interpretation, discourse relations, and commonsense entailment. *Linguistics and Philosophy*, **16**, 437–493.

Asher, N. & Lascarides, A. (2003). *Logics of Conversation*. Cambridge University Press.

Beaver, D. (2004). The optimization of discourse anaphora. *Linguistics and Philosophy*, **27**(1), 3–56.

Beaver, D. & Lee, H. (2003). Input-Output Mismatches in Optimality theory. In R. Blutner & H. Zeevat, editors, *Optimality Theory and Pragmatics*, pages 112–154. Palgrave, Basingstoke and New York.

Blutner, R. (2001). Some aspects of optimality in natural language interpretation. *Journal of Semantics*, **17**(3), 189–216.

Blutner, R. & Jäger, G. (1999). Competition and interpretation: The German adverbs of repetition.

Dale, R. & Reiter, E. (1996). The role of the gricean maxims in the generation of referring expressions. In *Proc. of the 1996 AAAI Spring Symposium on Computational Models of Conversational Implicature*. Stanford University, California, USA.

Gallese, V. (2003). A neuroscientific grasp of concepts: From control to representation. *Phil. Trans. Royal Soc.*, **358**, 1231–1240.

Gazdar, G. (1979). Pragmatics: *Implicature, Presupposition and Logical Form*. Academic Press, New York.

Grice, H. (1957). Meaning. *Philosophical Review*, **67**, 377–388.

Grosz, B. & Sidner, C. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, **12**(3), 175–204.

Heim, I. (1983). On the projection problem for presuppositions. In M. Barlow, D. Flickinger, & M.Westcoat, editors, *Second Annual West Coast Conference on Formal Linguistics*, pages 114–126. Stanford University.

Hendriks, P. & de Hoop, H. (2001). Optimality theoretic semantics. *Linguistics and Philosophy*, **24**, 1–32.

Hobbs, J. (1979). Coherence and coreference. *Cognitive Science*, **3**, 67–90.

Hobbs, J., Stickel, M., Appelt, D., & Martin, P. (1990). Interpretation as abduction. Technical Report 499, SRI International, Menlo Park, California.

Jǻger, G. (2000). Some notes on the formal properties of bidirectional optimality theory. In R. Blutner & G. Jager, editors, *Studies in Optimality Theory*, pages 41–63. University of Potsdam.

Jǻger, G. (2003). Learning Constraint Subhierarchies. the Bidirectional Gradual Learning Algorithm. In R. Blutner & H. Zeevat, editors, *Optimality Theory and Pragmatics*, pages 251–288. Palgrave, Basingstoke and New York.

Jasinskaja, E. (2007). *Pragmatics and Prosody of Implicit Discourse Relations: The Case of Restatement*. Ph.D. thesis, University of Tbingen.

Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, & H. Wettstein, editors, *Themes from Kaplan*, volume 135, pages 481–566. Oxford University Press, New York.

Kehler, A. (2002). *Coherence, Reference, and the Theory of Grammar*. CSLI Publications.

Mann, W. & Thompson, S. (1985). Rhetorical structure theory: Toward a functional theory of text organization. *TEXT Journal*, **8**, 243–281.

Mattausch, J. (2001). On optimization in discourse generation. ILLC report MoL-2001-04, MsC Thesis, University of Amsterdam.

Montague, R. (1974). Pragmatics. In R. Thomason, editor, *Formal Philosophy: Selected Papers of Richard Montague*, pages 95–118. Yale University Press, New Haven.

Polanyi, L. (1985). A theory of discourse structure and discourse coherence. In *Papers from the General Session at the Twenty-first Regional Meeting of the Chicago Linguist Society*, pages 25–27.

Prüst, H., Scha, R., & van den Berg, M. (1994). Discourse grammar and verb phrase anaphora. *Linguistics and Philosophy*, **17**, 261–327.

Reiter, E. & Dale, R. (2000). *Building Natural-Language Generation Systems*. Cambridge University Press, Cambridge.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, **1**, 75–116.

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry*, **27**, 720–731.

Sperber, D. & Wilson, D. (1984). *Relevance: Communication and Cognition*. Basil Blackwell, Oxford.

Umbach, C. (2001). Contrast and contrastive topic. In *Proceedings of the ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*. Helsinki.

Vallduvi, E. (1992). *The Informational Component*. Garland, New York.

van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, **9**, 333–377.

Van Rooy, R. (2003). Relevance and bidirectional OT. In R. Blutner & H. Zeevat, editors, *Pragmatics and Optimality Theory*, pages 173–210. Palgrave.

Zeevat, H. (2001). The asymmetry of optimality theoretic syntax and semantics. *Journal of Semantics*, **17**, 243–262.

Zeevat, H. (2006a). Applying an exhaustivity operator in update semantics. In M. Aloni, A. Butler, & P. Dekker, editors, *Questions in Dynamic Semantics*. Elsevier. originally appeared in 1994, H. Kamp, editor, Ellipsis, Tense and Questions, DYANA project deliverable.

Zeevat, H. (2006b). Freezing and marking. *Linguistics*, **44–5**, 1097–1111.

Zeevat, H. (2007). A full solution to the projection problem for presuppositions. ms, University of Amsterdam.

Zeevat, H. (2008). Constructive optimality theoretic syntax. In J. Villadsen & H. Christiansen, editors, *Constraints and Language Processing*, pages 76–88, ESSLLI Hamburg University.

Zeevat, H. (2009). Optimal interpretation as an alternative to Gricean pragmatics. In B. Behrens & C. Fabricius-Hansen, editors, *Structuring information in discourse: the explicit/implicit dimension*, Oslo Studies in Language. OSLA, Oslo.

Modelling discourse relations by topics and implicatures

The elaboration default

Ekaterina Jasinskaja
IMS Stuttgart/University of Heidelberg

This paper develops a theoretical approach that derives the semantic effects of discourse relations from the general pragmatic default priciples of exhaustivity—a kind of Gricean Quantity implicature—and topic continuity. In particular, these defaults lead to the inference of relations such as *Elaboration*, while other discourse relations, e.g. *Narration* and *List* are predicted to be 'non-default' and must be signalled, which contrasts with common assumptions in discourse theory. The present paper discusses some observations on the use of connectives and intonation in spontaneous speech which suggest that at least intonational signalling of such relations is obligatory.

This paper presents a theoretical approach to the inference of discourse relations. Previous work on this topic such as Hobbs (1985), Mann and Thompson (1988) Asher (1993), Kehler (2002) postulated a variety of inventories and taxonomies of discourse relations, which differed depending on the goals, empirical domains and philosophical assumptions of the study. These inventories provide a handy terminology and an insightful classification of facts about discourse. However, turning such an inventory into a basic theoretical construct always raises a range of hard questions: What is the right inventory? Why these relations and not others? etc.

The programmatic goal pursued in this study is to develop a theory of discourse relations that does not have to commit to a particular inventory, but derives the semantic and pragmatic effects associated with specific discourse relations from more general, independently motivated pragmatic principles. A lot of effort in developing an account of this kind has been made within the framework of Relevance Theory (see e.g. Blakemore, 2002). I will use some of the insights from this body of work, but will put emphasis on working out concepts that can be interpreted within the formal model-theoretic approach to meaning. The central role will be played by (a) the notion of discourse topic as the question under discussion (QUD), along the lines of Klein and von Stutterheim (1987), van Kuppevelt (1995), Ginzburg (1996),

Roberts (1996), Büring (2003), and (b) the mechanism of Gricean conversational implicature, represented here by one kind of Quantity implicature—exhaustive interpretation. Exhaustive interpretation of an answer *P* to a question *Q* says that *P* is the only (relevant) thing that has property *Q*, e.g. *Who snores? Bill.*—Bill is the only relevant individual who snores. It will be shown how a combination of exhaustivity as a default mode of interpretation and a principle of *topic continuity*, which bids you to stick to the topic and not to change it without warning, leads to the inference of such relations as *Elaboration*, i.e. relations that involve coreference between eventualities presented by adjacent sentences.

Event coreference relations will be so far the only group of discourse relations that will be modelled in the present paper; other relations will remain a task for the future. However, the theory I am going to propose will make one rather radical claim that has consequences for all kinds of relations. Since event coreference in the present approach results from applying very general default principles, it follows that relations like *Elaboration* are the default discourse relations. Hence, relations like *List*, *Narration*, or *Contrast* are not default and must be explicitly signalled. This goes, for instance, against the widely spread assumption that *Narration* is a default relation (cf. e.g. early Segmented Discourse Representation Theory, SDRT, Lascarides & Asher, 1993).[1] However, I will present a set of observations concerning the usage of discourse markers and intonation in spontaneous speech that provide some preliminary support for the marked character of *List* and *Narration* type relations.

This paper is structured as follows. Section 1 presents the philosophical motivation for the central role of the notions of QUD and exhaustive interpretation in the inference of discourse relations; Section 2 spells out the main positions of the theory, while Section 3 considers some relevant predictions. Finally, Section 4 discusses findings of some previous empirical studies on spontaneous dialogue that support the proposed approach.

1. Motivation

The focus of this study is on discourse relations that involve coreference relations between eventualities presented by the sentences, e.g. certain cases of *Elaboration* (1) and (2), and causal *Explanation* (3), cf. Danlos (2001), in contrast to relations that

1. It should be noted that the SDRT view of *Narration* as default developed as a reaction to an even more radical view, according to which the relation of temporal succession was not a default but part of the semantics of tense (Hinrichs, 1986; Kamp & Reyle, 1993). Thus we are in a way continuing the line started by Lascarides and Asher (1993), but go even further in relativising the role of *Narration* in discourse.

do not involve such coreference, e.g. *Narration* (4). The eventualities in (1) corefer in the sense that they describe the same action of Fred, in (2) it is the same event happening to Alena. According to Danlos (2001), the causal relation between the sentences in (3) can be derived by establishing coreference between hitting and the action part of breaking which in turn causes the subsequent "broken" state of the object. *Narration*, i.e. the relation of temporal sequence, is incompatible with event coreference since the latter implies simultaneity.

(1) Fred damaged a garment.
He stained a shirt.

(2) Alena broke her skis.
She lost her main means of transport.

(3) Fred broke the carafe.
He hit it against the sink.

(4) The lone ranger jumped on the horse
and (he) rode into the sunset.

A comprehensive explanatory theory of discourse interpretation must provide an answer to the questions how and why such relations are inferred, especially in cases where they are not signalled by any explicit markers like *because* (for *Explanation*) or *then* (for *Narration*), as in the examples above.

Previous approaches to the inference of discourse relations can be very crudely divided into two major groups: *coherence-based*, and *relevance-based* approaches. In the first group, the most basic assumption is that the discourse must be coherent, i.e. all sentences in a discourse must be connected by discourse relations from a designated set; these relations are thus inferred along with figuring out in which way a discourse fulfils the coherence requirement. One of the most prominent representatives of this view is SDRT (Asher & Lascarides, 2003). The other position does not view coherence as an aim in itself. Instead, a discourse must be *relevant*, i.e. fulfil its communicative goal in the situation in which it occurs. In this type of framework, coherence (and with it the inference of discourse relations) must be construed as a by-product of figuring out in which way a discourse is relevant. Pragmatic theories that are based directly or indirectly on Gricean ideas, e.g. Neo-Gricean pragmatics,[2] as well as approaches based on intentional structure such as Grosz and Sidner (1986) and the QUD-based models (Klein & Stutterheim, 1987; van Kuppevelt, 1995; Ginzburg, 1996; Roberts, 1996; Büring, 2003) can be counted to this category.

2. Relevance Theory can be seen as another instance, but it also stands apart since the RT notion of relevance has both a communicative and a cognitive component to it.

A relevance-based approach to discourse relations is appealing since it attempts to explain coherence rather than simply postulating it, but this issue is nevertheless controversial. One of the most challenging points of criticism put forward by Asher and Lascarides (2003) is that making the inference of discourse relations entirely dependent on the recognition of the speakers' goals or intentions as well as any other "private" features of their mental states introduces unnecessary conceptual and computational complexity into the model. Not always, but often discourse relations can be successfully inferred by the hearer without having perfect information about the speaker's communicative intentions. The following example illustrates this.

Suppose *A* sees *B* all black and blue and eyes swollen with tears; *A* asks *What happened?*; speaker *B* gives the answer in (5), and *A* eventually notices pieces of broken glass on the floor.

(5) **A:** What happened?
B: Fred broke the carafe.
He hit it against the sink.

Hearing this answer, *A* will probably have doubts whether *B*, unintentionally or deliberately, got his question right. The question *A* had meant was 'What happened to you that made you weep and caused the bruises?' The question answered by *B* is apparently 'What happened such that there is all this broken glass on the floor?' But in spite of this "misunderstanding," *A* will be able to infer from *B*'s answer that Fred broke the carafe *by* hitting it against the sink. That is, the inference of coreference between the breaking and the hitting event in (5) does not require sharing the communicative goals by the speakers.

The main purpose of the present study is to formulate a fragment of a relevance-based theory of discourse relations which nevertheless can accommodate the above observation. Rather than developing special machinery for dealing with coherence, semantic effects of discourse relations should be derived from relations between discourse goals associated with the sentences. However, the theory should *predict* where and when the exact knowledge of the underlying goals is necessary or unnecessary for the inference of a discourse relation. Thus, the fact that this information is unnecessary in (5) should be derived as a theorem in this framework.

In this paper, communicative goals of utterances are modelled as questions under discussion. Following Schulz and van Rooij (2006), the exhaustive interpretation of an utterance with respect to its QUD is intended to implement the Gricean mechanism of conversational implicature—the pragmatic meaning that comes on top of the conventional semantics of the sentence which results from the assumption of the speakers' rational and cooperative communicative behaviour.

2. Outline of the theory

The present proposal is cast in the framework of dynamic update semantics (MDPL, Dekker, 1993) enriched with a notion of exhaustive update (Zeevat, 1994; Schulz & van Rooij, 2006). In this framework, the meaning of a sentence represents its context change potential. It is defined as an *update* function ($[\![\varphi]\!]$) that takes the initial context, or *information state* (s), as an argument and returns a new information state ($s\,[\![\varphi]\!]$) as its value. The information state represents the current common ground of the discourse participants and it is a set of world-assignment pairs: it contains only worlds consistent with the information in the common ground and only the assignment functions that assign appropriate objects of the domain to variables (e.g. pronouns) introduced so far in the discourse. The non-exhaustive update function corresponds to the literal meaning of a sentence, whereas the exhaustive update to its exhaustive interpretation. The interpretation of a discourse—a monologue or a dialogue turn of a single speaker produced without intervention from other discourse participants—in turn is a sequence of exhaustive and non-exhaustive updates of the initial information state s with the meanings of individual utterances (represented schematically in Figure 1). As Figure 1 is intended to suggest, the update function is sensitive to the QUD or the *discourse topic* T of the current utterance. The goal of this section is basically to explain Figure 1. Section 2.1 cites the necessary definitions that elucidate the relationship between the topic and the (exhaustive) interpretation of an utterance. Then Section 2.2 presents some constraints on topics and other parameters of discourse update.

2.1 Definitions

The notion of QUD or discourse topic can be implemented formally in a number of different ways. In this paper it will be identified with what is often called the *question predicate* or the *question abstract*—an atomic predicate symbol or a

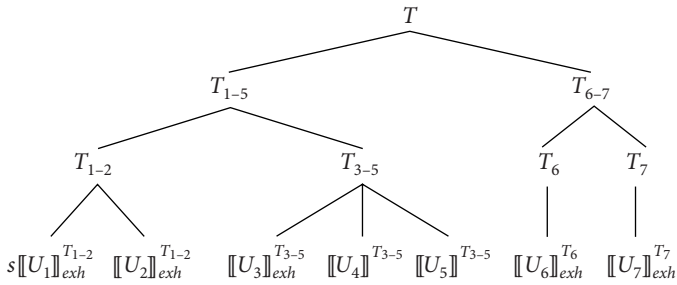


Figure 1. Interpretation of a sequence of utterances $\langle U_1, \ldots, U_7 \rangle$ and its topic structure

complex λ-term that is obtained by abstracting over the *wh*-elements of an interrogative sentence, e.g. the predicate *happen* for the question *What happened?*, or λ*x* [*kissed(john, x)*] for *Who did John kiss?* The non-exhaustive update $s[\![\phi]\!]^T$ wrt. the predicate *T* is defined just like the standard dynamic update function with an additional definedness condition that *T* be contained in ϕ.[3]

For the exhaustive update, I borrow the definition of dynamic exhaustification (6) proposed by Schulz and van Rooij (2006):

(6) $s[\![\phi]\!]^T_{exh} = min_{<T}(s[\![\phi]\!]^T)$

The exhaustive update $s[\![\phi]\!]^T_{exh}$ is a subset of the non-exhaustive update $s[\![\phi]\!]^T$ that only contains world-assignment pairs that are minimal wrt. a topic-sensitive order $<_T$.[4] The latter orders world-assignment pairs along the inclusion relation between the extensions of the predicate *T* in their respective worlds while keeping other properties of those worlds as well as the assignment functions constant:[5]

(7) $\langle w_1, g_1\rangle <_T \langle w_2, g_2\rangle$ iff
  a. $F(T)(w_1) \subset F(T)(w_2)$, and
  b. for every *R* that is independent from *T*: $F(R)(w_1) = F(R)(w_2)$, and
  c. $g_1 = g_2$

Thus minimisation wrt. $<_T$ boils down essentially to minimisation of the the extension of the question predicate *T*, which in the case of $s[\![T(a)]\!]_{exh}$, for instance, leads to the implicature that e.g. *a* is the *only* individual that has property *T*.

---

3. Note that the natural language sentence that corresponds to φ might not and often will not overtly contain *T*. The predicate will have to be recovered at the level of semantic representation, perhaps in a similar way as elided material is recovered.

4. For an adequate account of event coreference relations embedded in longer discourses one might need to adjust the definition of exhaustive update to disregard the information on *T* accumulated in the common ground when minimising the extension of *T*. This is achieved in (i) by applying minimisation to the non-exhaustive update of the "minimal" information state s_0 that contains no information about the world except some lexical semantic restrictions (meaning postulates) and highly conventionalised world knowledge, but preserves all information about the introduced variables.

(i) $s[\![\phi]\!]^T_{exh} = s[\![\phi]\!]^T \cap min_{<T}(s_0[\![\phi]\!]^T)$

For space reasons the discussion of this modification is skipped here, but see Jasinskaja (2007) for details.

5. The appropriate restriction (cf. the *independence* requirement) in condition (7b) is notoriously difficult to define. See Jasinskaja (2007, pp. 217–218) for a brief discussion.

2.2 Constraints

The desired theory of discourse interpretation is expected to provide the parameters on which discourse update depends: the segmentation of the discourse flow into basic update units, utterances (e.g. U_1, U_2, etc. in Figure 1), the choice of topic T, as well as the choice between the exhaustive $[\![\cdot]\!]_{exh}$ and the non-exhaustive update $[\![\cdot]\!]$.

As a first approximation let's assume that an utterance equals a sentence. Determining the topic is a particularly difficult task since topics, as reflections of the agents' discourse goals, belong to the set of *a priori* "private features" of the agents' mental states that Asher and Lascarides (2003) talked about. However, one can already get quite far by adopting a number of constraints that regulate relations between topics without determining the topics themselves. A constraint that is of special interest for our present purposes is the *Principle of Topic Continuity*, named so after Givón (1983) and first applied to QUDs by Zeevat (2005). The principle requires that by default subsequent utterances be interpreted with respect to the same question predicate T:[6]

(8) *The Principle of Topic Continuity:*
By default, the discourse topic does not change.

Concerning the choice between exhaustive and non-exhaustive update, the constraint in (9) states that by default, update is exhaustive.[7]

(9) *The Principle of Exhaustive Interpretation:*
By default, an utterance is interpreted exhaustively.

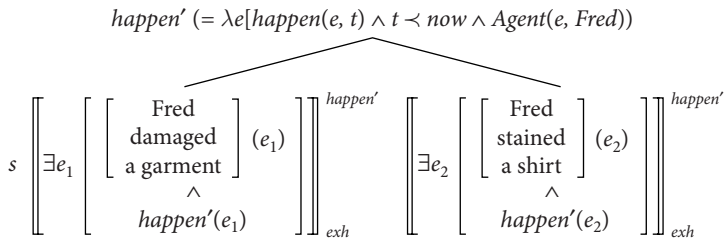


Figure 2. Interpretation of (1)

6. In Figure 1 the sequences of utterances $\langle U_1, U_2\rangle$ and $\langle U_3, U_4, U_5\rangle$ satisfy topic continuity, whereas e.g. in $\langle U_6, U_7\rangle$ this principle is violated at the local level.

7. This principle is violated in utterances U_4 and U_5 in Figure 1.

Both (8) and (9) are default principles and a crucial question is under which circumstances they can be overridden. In the rest of this paper I explore the consequences of assigning a relatively high rank to these constraints. It will be assumed that these defaults can only be cancelled by explicit linguistic cues that either indicate a specific discourse relation (e.g. *then*), a topic change (e.g. contrastive topic marking), or non-exhaustivity (e.g. continuation intonation).

3. Predictions

3.1 The default case

The discourse in (1) presents a case where both default principles apply, since the sentences do not contain any explicit markers of discourse relations, topic change, etc. The sentences are connected asyndetically and we will assume that they are uttered with the neutral (falling) declarative intonation, indicated by a period at the end of each sentence. Thus both sentences are interpreted exhaustively with respect to the same topic. Let's assume that the topic is the question predicate λe [*happen* (*e,t*) $\wedge$ $t \prec$ *now* $\wedge$ *Agent* (*e, Fred*)], i.e. *what did Fred do?* or what happened such that Fred was the agent of that event, abbreviated as *happen*' in Figure 2.

For simplicity, suppose that the initial information state *s* contains no information on the actions of Fred, so it will equally contain worlds where there is one, two, three, etc. events in the extension of *happen′*, where those events belong to one, two, three or more different types, e.g. doing the dishes, having a row with the neighbour, as well as Fred damaging a garment. The non-exhaustive update of *s* with the first sentence *Fred damaged a garment* will only contain world-assignment pairs where there is at least one event of Fred damaging a garment in the extension of *happen′* and the referent of e_1 is mapped to that event. Since all the worlds where *F*(*happen′*) contains more than just that event are *happen′*-greater (in terms of $<_{happen'}$, cf. (7)) than the worlds where Fred damaged a garment exactly once and nothing else *happen′*-ed, only the latter will survive in the exhaustive update, i.e. after minimisation wrt. $<_{happen'}$, cf. (6). Thus in all worlds of the resulting information state the *extension of happen*' will contain a single event that matches the description *Fred damaged a garment*. The subsequent update with the second sentence will only pick out those worlds where that event also matches the description *He stained a shirt & happen′*, since there simply is no other event that this description could consistently apply to. As a result, the information state will entail that the two event types are always instantiated by the same event individual, so an event coreference relation is derived between the sentences. In other words, Fred damaged a garment and this is the only thing that he did, and Fred stained a shirt and this is the only thing that he did. Therefore these events *must* be the same.

Note that the choice of question predicate does not affect this inference. Suppose that the topic in (5) is the question predicate *happen′* = λ*e* [*happen* (*e,t*) ∧ *t* ≺ *now* ∧ *cause*(*e, bruises*)], i.e. what happened that caused *B*'s bruises. The coreference between breaking the carafe and hitting it against the sink is inferred in the same way, by identity of the only event that *happen′*-ed with Fred breaking the carafe on the one hand and Fred hitting it against the sink on the other. If *A*'s question in (5) is interpreted as 'What happened such that there is all this broken glass on the floor?,' i.e. if the question predicate is *happen″* = λ*e* [*happen (e, t)* ∧ *t* ≺ *now* ∧ *cause*(*e, broken glass*)], the result is the same: since Fred breaking the carafe is the only event that *happen″*-ed and his hitting it against the sink is also the only event that *happen″*-ed, they must be the same event. If the sentences were interpreted with respect to distinct topics, event coreference would not necessarily follow. However, as long as the topic is the same, which is a consequence of the principle of topic continuity (8), the inference is valid *regardless* of what exactly the topic is. Thus the proposed relevance-based theory predicts that perfect information of the participants' communicative goals is unnecessary for the inference of discourse relations based on event coreference, and therefore meets Asher and Lascarides' critique (discussed in Section 1), at least for this group of discourse relations.

The present approach also predicts that event coreference is inferred always if both default principles—topic continuity and exhaustive interpretation—are maintained. This means that at least one of the defaults must be violated in order to establish some other discourse relation, e.g. *Narration*, which involves temporal succession of events and is therefore incompatible with event coreference. Alternatively, a discourse relation must be established utterance-internally (within the scope of a single exhaustification operator). The next section discusses the role played by continuation intonation and the conjunction *and* in this process.

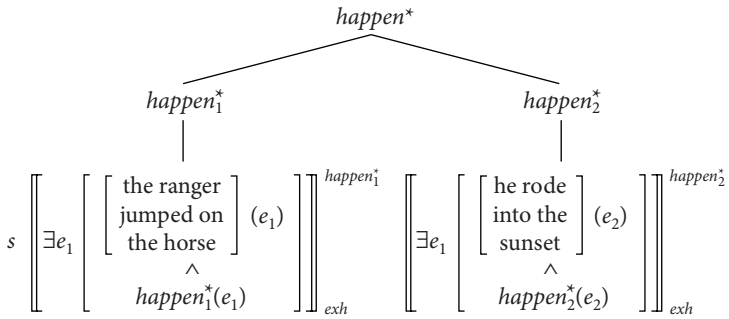


Figure 3. Interpretation of (11)

3.2 The effect of continuation intonation and the conjunction *and*

As was already mentioned, this paper explores the possibility that the principles of topic continuity and exhaustive interpretation can only be violated by explicit linguistic means that either encode (a) a specific discourse relation, e.g. *then* for *Narration*, (b) a topic change, e.g. contrastive topic accent, or (c) non-exhaustivity, e.g. utterance-final continuation intonation. This section will concentrate on continuation intonation and the conjunction *and*.

There are two possibilities concerning the analysis of *and*. One possibility is to assume after Blakemore and Carston (1999) that the conjunction makes one sentence/utterance out of two, i.e. conjoined clauses constitute a single update unit. Thus in (4) neither default principle is violated; however, both clauses are interpreted in the scope of a single exhaustification operator:

(10) $s[\![[\exists e_1[\text{Ranger jumped on horse}(e_1) \wedge \textit{happen}^\star(e_1)] \wedge$
$\exists e_2[\text{Ranger rode into sunset}(e_2) \wedge \textit{happen}^\star(e_2)]]\!]^{\textit{happen}^\star}_{exh}$

The resulting information state entails that an event of type the lone ranger jumping on the horse happened in some relevant sense *happen**; in addition, his riding into the sunset *happen**-ed, but nothing else. It does not follow from this that jumping on the horse and riding must be the same event. In fact, the world knowledge will probably exclude the coreference reading and create a bias for temporal succession.

The second possibility for the analysis of *and*, which might be necessary to account for cases of sentence-initial *and* like (11), is to view the conjunction as a topic management device that indicates that the current topic (question predicate) T_2 is different from the previous topic T_1 but stands in an *additive relation* to it, i.e. there is an overarching topic T such that $T = T_1 + T_2$, where (12) could be a first approximation for the definition of +.[8]

(11) The lone ranger jumped on the horse.
And he rode into the sunset.

(12) For predicate symbols T, T_1, and T_2:
$T = T_1 + T_2$ iff
$\forall w[F(T)(w) = F(T_1)(w) \cup F(T_2)(w) \wedge F(T_1)(w) \cap F(T_2)(w) = \emptyset]$

The definition says that in every world the extension of T_1 and T_2 partitions the extension of T. For example, suppose *happen** is understood as what happened at the end of the film, and suppose the time span meant by "end of the film" can

8. The analysis of *and* as an additive marker is worked out in Zeevat and Jasinskaja (2007) and Jasinskaja and Zeevat (2009).

be partitioned into subsets t_1 and t_2 of time points. Then *happen** can constitute the sum of *happen*$^\star{}_1$ 'happen at a some point in t_1' and *happen*$^\star{}_2$ 'happen at some point in t_2' in the sense of (12), so *happen*$^\star{}_1$ and *happen*$^\star{}_2$ form a possible topic sequence triggered by sentence-initial *and*. Thus (11) can be interpreted as shown in Figure 3. The exhaustive interpretation principle is maintained, but topic continuity is violated. The sentences are exhaustivized wrt. distinct topics, so the event coreference effect is blocked: if the lone ranger jumping on the horse is the only event that happened at t_1 and his riding into the sunset is the only event that happened at t_2, it does not follow that riding and jumping are the same event.

Finally, let's consider continuation intonation (↗), which is typically realised as an utterance final pitch rise distinct from completion intonation in a vast variety of languages (Cruttenden, 1981). Some uses of (↗) are probably best analysed along the same lines as sentence-internal *and*, cf. (10), as indicating that the utterance, and hence the update unit, is not complete yet, so exhaustive update should be postponed until the next completion. However, here I will concentrate on intonational realisation of items in *open lists* like (13):

(13) The lone ranger jumped on the horse (↗) ,
he rode into the sunset (↗) …

Since a completion might never come, such a discourse does not provide a unit to interpret exhaustively in any reasonable sense; nevertheless, the hearers seem to have no difficulty in understanding such "incomplete" discourses. Therefore, I propose that open list intonation signals that the current utterance should undergo non-exhaustive update:

(14) $s[\![\exists e_1[\text{Ranger jumped on horse}(e_1) \wedge \mathit{happen}^\star\,(e_1)]]\!]^{\mathit{happen}^\star}$
$[\![\exists e_2[\text{Ranger rode into sunset}(e_2) \wedge \mathit{happen}^\star\,(e_2)]]\!]^{\mathit{happen}^\star}$

This analysis represents (13) as a usual dynamic conjunction of the sentences. It says nothing about whether the listed events exhaust the relevant *happen**-ings, and it has no entailments regarding coreference relations between e_1 and e_2. This illustrates the case where the event coreference effect is cancelled by violating exhaustivity, while maintaining topic continuity.

In sum, this section has shown three different ways of cancelling the event coreference effect that arises in the default case. This cancellation is necessary in order to make possible the inference of discourse relations such as *Narration*, *List*, *Parallel*, etc., which either have incompatible or weaker semantics than event coreference. Obviously, additional machinery is needed to actually *infer* e.g. temporal relations in the case of *Narration*, here we were only concerned with consistency of such an inference with the rest of the theory. Crucially however, given the assumption that the default principles of exhaustive interpretation and

topic continuity can only be violated by explicit linguistic cues, the proposed approach implies that *Narration* and other relations in its category must be explicitly triggered, at least by such semantically weak cues as the conjunction *and* or continuation intonation. The next section discusses the empirical validity of this implication.

4. Intonation and conjunction in spontaneous speech

The approach presented above leads to a rather strong prediction that discourse relations such as *Narration* or *List* cannot be inferred between two asyndetically connected sentences that are both uttered with (falling) completion intonation. At first glance this appears wrong. The literature on pragmatics abounds in examples like (15) where it is claimed that the sentences present a temporal sequence of events, although they are not connected by a conjunction *and*, and although they both end in a period, which normally corresponds to a final fall in reading.

(15) The lone ranger jumped on the horse.
He rode into the sunset.

However, there are reasons to believe that spontaneous speech is different from written language and read speech in this respect. This section recapitulates some previous findings that suggest that asyndetic connection in combination with completion intonation is indeed somewhat suboptimal with *Narration*.

First of all, it is an old observation that the sentence-initial conjunction *and* is much more frequent in spontaneous spoken narrative than in writing (see e.g. Chafe, 1982). In spoken narrative almost every utterance often starts with an *and*, as in the discourse in (16) cited by Schiffrin (1987).

(16) **A:** a. You lived in West Philly? Whereabouts?
B: b. Well, I was born at 52nd and em… tsk… oh: I forgo- well……
I think its 52nd and Chew.
c. *And* um… *and* uh I grew up really in the section called Logan.
d. *And* then, I went into the service, for the two years,
e. *and* then when I came back, I married… I- I- I got married.
f. *And* I- then I lived at uh 49th and Blair.

For example, if the conjunctions were removed in (16), and especially if the *Narration* markers *then* were removed as well, the discourse would sound much less natural, or at least much less "conversational" with the utterance-final falling completion intonation, indicated by periods in the conversation transcripts.

At the same time, it is not the case that asyndetic connection does not co-occur with completion intonation at all in spontaneous speech—only when

they co-occur, the discourse relation is normally *Elaboration* or *Explanation*, rather than *Narration*, as in the following examples that illustrate *Elaboration*:

(17) **A:** a. Do either one of your daughter in laws work?
B: b. No but they did.
c. Both my daughters in laws worked.

(18) a. *And* uh: that's- that's the answer.
b. That's why I say they're the most prejudiced.

At least part of the reason for this distribution could be a more restricted usage of asyndetic connection with *Narration* in spoken language. In order to prevent an event coreference interpretation, the speakers have to use an *and* after completion intonation either to put the next clause into the scope of the same exhaustification operator as the previous one, or to indicate that the topic of the current sentence stands in an additive relation to the previous topic: what happened? what else happened? etc., cf. (12). The following example from a cartoon retelling cited in Carroll et al. (2008) illustrates this most clearly. Whenever the utterance presents the next event in the sequence a connective *and* or *so* is used, as in (19c), (19e), and (19g), whereas the asyndetically connected utterances (19d) and (19f) redescribe or elaborate on the previously introduced events.

(19) a. …
b. *and* he goes over to it
c. *and* he kneels
d. gets down on his hands and his knees in front of it
e. *and* he feels the paper
f. he feels that it's wet
g. *so* he reaches his hands up to wait for the next drop of water
h. …

Further support for our hypothesis comes from Nakajima and Allen's (1993) study of intonation in elicited spontaneous task-oriented dialogue.[9] For the purposes of the study, the dialogue was segmented into utterance units (UUs) and the transitions between consecutive utterance units were annotated with discourse relations, which included what Nakajima and Allen called *elaboration class* and *speech-act continuation* transitions. The *elaboration class* relations hold when "the current

9. The participants of the study are involved in a game, where the task of one participant, called "Human" (H) is to achieve a specific goal by making plans to manufacture and ship various goods to specified locations in the game's world by the due date. The other participant, called "System" (S), has up-to-date information on the state of the world and assists H in making plans to achieve the given goal. The authors obtain about three hours of spontaneous dialogue.

utterance adds some relevant information to the previous utterance." Judging by the examples given in the paper, cf. (20) and (21), this transition type corresponds roughly to our notion of *Elaboration* based on event coreference.

(20) **H:** a. are there oranges available in warehouses in both cities H and I
S: b. uhh let's see
there're oranges available in uhh yes, in H and in city I
c. They have oranges in both places, enough for uhh uhm
several boxcars of oranges

(21) **H:** a. let's do that
b. let's move E2 to city E

Speech-act continuation holds when "a single speech-act continues over several UUs." The authors note that most of the speech-act continuations occur in "sequential conjunctions," cf. (22). Note that speaker H presents a list of actions that should be taken one after the other, thus in our present terms, the discourse relation between (22c) and (22d) could be analysed as *Narration*.

(22) **H:** a. now let's uhh
assume the oranges are already loaded into the boxcar B6
S: b. hnn-hnn
H: c. and we'll take the engine that's at city H
d. we'll move the boxcar with engine down to city A

Among other prosodic features, Nakajima and Allen study utterance-final fundamental frequency (F_0). They find that F_0 at the end of an utterance preceding an *elaboration class* boundary tends to be significantly *lower* than before a *speech-act continuation* boundary. In other words, utterances like (20b) and (21a) are more likely to be pronounced with a lower final *F*, presumably associated with a completion fall, whereas utterances like e.g. (22c) end high more often, which could be a reflex of various kinds of continuation tunes. This is again consistent with the prediction that the usage of completion intonation is not typical for relations like *Narration*.

But if our predictions are born out in spoken language, the question that arises immediately is why written language is different in that it allows narrative sequences to be expressed by asyndetically connected sentences ending with a period, like (15). One possible answer is that the function of period in writing is dramatically different from that of completion fall in speaking. However, this hypothesis is not so easily substantiated given the pervasive tendency to read periods as falls. Another possibility is that the role of the opposition of asyndetic connection and the sentence-initial conjunction *and* is not the same in the two speech modes. Historical studies on the development of written language support

this hypothesis. For instance, Dorgeloh (2004) finds that sentence-initial *and* used to be much more frequent in texts of the Early Modern English period. Sentence-initial *and* in texts of that period is particularly typical of narrative sequences, and the narrative sequence in turn constitutes the prevailing discourse strategy not only in texts narrative "by nature," such as historical texts, but also in scientific prose where evidence is often recounted in the form of experience. Dorgeloh's corpus study shows a clear decline of the frequency of sentence-initial *and* in both scientific and historical texts, especially between the first and the second stage of the Early Modern English period (1500–1570 vs. 1570–1640). Dorgeloh explains this decline by a general shift from narrative to argumentative organisation of scientific text. Translating this into the terminology of discourse relations, scientific writing experienced increasing avoidance of *Narration*, presumably, in favour of such discourse relations as *Explanation* and *Evidence*. Thus "the usage of sentence-initial *and* became associated with the older, more narrative, and hence less professional style and thus became increasingly stigmatised" (Dorgeloh, 2004, p. 1770) which ultimately led to its more general banishment from larger parts of the written language, even from narratives. This must have affected the division of labour between sentence-initial *and* and asyndetic connection, expanding the usage of the latter to include *Narration*. In an optimality-theoretic setting, it would be elegant to model this property of written language by introducing a relatively high-ranked stylistic constraint "Avoid sentence-initial *and*," which is absent in the discourse model of spontaneous dialogue. The development of this proposal goes beyond the scope of this paper.

Finally, it should be noted that previous empirical studies have mainly concentrated either on discourse markers or on intonation in isolation, but the proposed theory makes predictions primarily on the interaction between the two. Therefore conclusive evidence for or against our proposal can only be gained when intonation and conjunction are studied simultaneously, which remains a task for the future.

5. Conclusions and outlook

This paper has presented a fragment of a theory of discourse relations that has the explanatory appeal of relevance-based approaches to pragmatics and at the same time meets some points of criticism raised by the opponents of such approaches. The central role played by the principles of topic continuity and exhaustive interpretation in this framework, which remained concealed by the specifics of the written mode of communication, revealed itself once we took a closer look at spoken language. There is no need to repeat that the present proposal has very

limited coverage, only providing an inference mechanism for the default relations involving event coreference, i.e. certain kinds of *Elaborations* and *Explanation*, and a possibility to cancel them by explicit marking. There remain a number of open issues some of which I will mention only briefly.

First of all, the very notion of event identity is rather controversial. There are views supporting more fine-grained (e.g. Dretske, 1977) and cruder identity notions (e.g. Davidson, 1967). The notion we presupposed here is obviously among the cruder ones, roughly, identifying events with the spatio-temporal chunks they occupy. However, even under this relatively weak assumption, there remains a lot of room for vagueness and ambiguity, which, however, are probably characteristic of any notion of identity, be that of eventualities or normal individuals. For instance, there are more central, more essential parts and more periferal parts to an object. In the case of a human body we would certainly consider the head, the chest, and e.g. the inner organs as part of it, but we might have doubts about hair, finger nails, or clothing (see Kratzer, 1989, p. 610, for related discussion). The take-off and the landing might be seen as more constitutive parts of a flight than e.g. the security measures demonstration or the check-in. This issue becomes relevant in examples like (23), discussed by Danlos (1999):

(23) Nicholas flew from Austin to Paris.
He took off at 6 a.m. and landed at 2 p.m.

Strictly speaking, a flight cannot be reduced to just the take-off and the landing, the process in between is obviously part of it, too. Therefore the complex event constituting the sum of the take-off and the landing forms a proper part of the flight. This lack of identity should make it impossible to connect the sentences asyndetically with the default completion intonation as in (23), but the discourse is nevertheless felicitous. Apparently, for licensing *Elaborations* like (23) it is enough if all the *essential* parts of the first event are named,[10] and what is essential is influenced by the communicative situation.

---

10. If the second sentence only names some periferal part of the event in the first sentence, then the discourse as it stands, without a continuation, is quite infelicitous with asyndetic connection and the falling completion intonation pattern:

(i) ??The council built the bridge (↘) .
They got an architect to draw up the plans (↘) . [end of the story]

Such a discourse can at best be understood as a joke meaning that drawing up plans is all the bridge building was about. It becomes better though, if the second sentence is uttered with a rising continuation intonation.

A more advanced version of the proposed theory should account for this observation. A possible mechanism for making event identification sensitive to contextual factors is built in within the present indirect way of dealing with event coreference via continuing topics. For instance, the topic predicate can be taken to be what Schulz and van Rooij (2006) call the *optimal* question predicate whose extension may deviate from its conventional extension on demands of relevance and, in particular, be made dependent on the degree of granularity with which the communication participants are looking upon the domain (see Jasinskaja, 2007, for more detailed discussion).[11] However, the details of this approach still need to be worked out.

The second area in which the proposed theory calls for extension concerns a whole range of other discourse relations which can be inferred by default, i.e. without special marking by connectives or intonation, but which do not involve event coreference, not even in the weak sense discussed above. This group includes all kinds of non-causal *Explanations* like (24), relations of *Evidence* like (25) and (26), and *Background* (27).

(24) John and Mary baptized all their children. They are good catholics.

(25) John must have been here recently. There are his footsteps.

(26) Max broke his leg. John told me.

(27) He walked into the room. The director was slumped in her chair.

A possible solution strategy could follow the idea mentioned by Zeevat (this volume), according to which a topic is only settled by a proposition plus all the evidence (sensory evidence, theoretical proof, or hearsay) that forms the speaker's knowledge of that proposition. Therefore it is not generally enough to say *John must have been here*—one also needs to convince the hearer (to a relevant extent) in order to settle the topic. In that sense sentences connected by *Explanation* and *Evidence* can also be viewed as pertaining to the same topic. In the case of *Background*, the speaker tries to provide "a full picture" of the described event, not only reporting the event itself, but also all sensory evidence around it. But obviously, these ideas also require further development.

In sum, there remain lots of open issues, both empirical and theoretical, and even once resolved, the present proposal only gives a rather basic division of discourse relations into two classes—"default" vs. "non-default." Nevertheless,

---

11. Another approach to event coreference by default identifies the event variables introduced by the verbs directly, by-passing topics and exhaustification, as is done e.g. by Zeevat (2006). The simplicity of this approach is appealing, but it is even less obvious how it could account for the context-sensitive nature of event identity.

I have tried to show that the proposed approach is promising and gains on explanatory elegance and simplicity, making the inference of discourse relations integral part of general pragmatics.

References

Asher, N. (1993). *Reference to Abstract Objects in Discourse.* Kluwer, Dordrecht.

Asher, N. & Lascarides, A. (2003). *Logics of Conversation.* Studies in Natural Language Processing. Cambridge University Press.

Blakemore, D. (2002). *Relevance and Linguistic Meaning: The semantics and pragmatics of discourse markers.* Cambridge University Press.

Blakemore, D. & Carston, R. (1999). The pragmatics of *and-conj*unctions: The non-narrative cases. *UCL Working Papers in Linguistics,* 11: 1–20.

Büring, D. (2003). On D-trees, beans and B-accents. *Linguistics and Philosophy,* 26: 511–545.

Carroll, M., Roßdeutscher, A., Lambert, M., & von Stutterheim, C. (2008). Subordination in narratives and macrostructural planning: Taking a comparative point of view. In Fabricius-Hansen, C. & Ramm, W., editors, *'Subordination' versus 'Coordination' in Sentence and Text: A cross-linguistic perspective,* Amsterdam. Benjamins. To appear.

Chafe, W. (1982). Integration and involvement in speaking, writing, and oral literature. In Tannen, D., editor, *Spoken and Written Language: Advances in Discourse Processes,* volume 9, pages 35–54. Ablex, Norwood.

Cruttenden, A. (1981). Falls and rises: Meanings and universals. *Journal of Linguistics,* 17: 77–91.

Danlos, L. (1999). Event coreference between two sentences. In *Proceedings of the 3rd Lnternational Workshop on Computational Semantics,* Tilburg.

Danlos, L. (2001). Event coreference in causal discourses. In Bouillon, P. & Busa, F., editors, *The Language of Word Meaning,* pages 216–241. Cambridge University Press.

Davidson, D. (1967). The logical form of action sentences. In *Essays on Actions and Events.* Oxford University Press, Oxford.

Dekker, P. (1993). *Transsentential Meditations.* Ph.D. thesis, ILLC-Department of Philosophy, University of Amsterdam.

Dorgeloh, H. (2004). Conjunction in sentence and discourse: Sentence-initial *and* and discourse structure. *Journal of Pragmatics,* 36: 1761–1779.

Dretske, F.I. (1977). Referring to events. *Midwest Studies in Philosophy,* 2: 90–99.

Ginzburg, J. (1996). Dynamics and the semantics of dialogue. In Seligman, J. & Westerståhl, D., editors, *Logic, Language and Computation,* volume 1. CSLI Publications, Stanford, CA.

Givón, T. (1983). Topic continuity in discourse: An introduction. In Givón, T., editor, *Topic Continuity in Discourse,* pages 1–42. John Benjamins, Amsterdam.

Grosz, B.J. & Sidner, C.L. (1986). Attention, intentions and the structure of discourse. *Computational Linguistics,* 12(3): 175–204.

Hinrichs, E. (1986). Temporal anaphora in discourses of English. *Linguistics and Philosophy,* 9: 63–82.

Hobbs, J.R. (1985). On the coherence and structure of discourse. Technical Report CSLI-85–37, Center for the Study of Language and Information, Stanford University.

Jasinskaja, E. (2007). *Pragmatics and Prosody of Implicit Discourse Relations: The Case of Restatement.* Ph.D. thesis, University of Tübingen.

Jasinskaja, K. & Zeevat, H. (2009). Explaining conjunction systems: Russian, English, German. In Riester, A. and Solstad, T., editors, *Proceedings of Sinn und Bedeutung 13,* volume 5 of *SinSpeC. Working Papers of the SFB 732.* University of Stuttgart.

Kamp, H. & Reyle, U. (1993). *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory.* Kluwer Academic Publishers.

Kehler, A. (2002). *Coherence, Reference, and the Theory of Grammar.* CSLI Publications.

Klein, W. & von Stutterheim, C. (1987). Quaestio und die referentielle Bewegung in Erzählungen. *Linguistische Berichte,* 109: 163–185.

Kratzer, A. (1989). An investigation of the lumps of thought. *Linguistics and Philosophy,* 12(5): 607–653.

Lascarides, A. & Asher, N. (1993). Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy,* 16: 437–493.

Mann, W.C. & Thompson, S. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text,* 8(3): 243–281.

Nakajima, S. & Allen, J.F. (1993). A study on prosody and discourse structure in cooperative dialogues. *Phonetica,* 50: 197–210.

Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *OSU Working Papers in Linguistics,* 49: 91–136.

Schiffrin, D. (1987). *Discourse Markers.* Cambridge University Press, Cambridge.

Schulz, K. & van Rooij, R. (2006). Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy,* 29(2): 205–250.

van Kuppevelt, J. (1995). Discourse structure, topicality and questioning. *Journal of Linguistics,* 31: 109–147.

Zeevat, H. (1994). Applying an exhaustivity operator in update semantics. In Kamp, H., editor, *Ellipsis, Tense and Questions,* pages 233–269. ILLC, Amsterdam. Dyana-2 deliverable R2.2.B.

Zeevat, H. (2005). Optimality theoretic pragmatics and rhetorical structure. Handout of the talk at the Sixth International Tbilisi Symposium on Language, Logic and Computation. September 12–16, 2005, Batumi.

Zeevat, H. (2006). Optimal interpretation as an alternative to Gricean pragmatics. In Fabricius-Hansen, C. & Ramm, W., editors, *Proceedings of the SPRIK Conference 2006: Explicit and Implicit Information in Text. Information Structure Across Languages.* To appear.

Zeevat, H. & Jasinskaja, K. (2007). *And* as an additive particle. In Aur-nague, M., Korta, K., & Larrazabal, J.M., editors, *Language, Representation and Reasoning. Memorial volume to Isabel Gomez Txurruka,* pages 315–340. University of the Basque Country Press.

The role of logical and generic document structure in relational discourse analysis

Maja Bärenfänger, Harald Lüngen,
Mirco Hilbert & Henning Lobin
Fachgebiet Angewandte Sprachwissenschaft und Computerlinguistik,
Institut für Germanistik, Justus-Liebig-Universität Gießen

This study examines what kind of cues and constraints for discourse interpretation can be derived from the logical and generic document structure of complex texts by the example of scientific journal articles. We performed statistical analysis on a corpus of scientific articles annotated on different annotations layers within the framework of XML-based multi-layer annotation. We introduce different discourse segment types that constrain the textual domains in which to identify rhetorical relation spans, and we show how a canonical sequence of text type structure categories is derived from the corpus annotations. Finally, we demonstrate how and which text type structure categories assigned to complex discourse segments of the type "block" statistically constrain the occurrence of rhetorical relation types.

1. Introduction

The theme of this article is a corpus-based investigation of the role of logical and generic document structure in the relational discourse analysis of complex texts by the example of scientific journal articles. One aim of this is to formulate cues and constraints such that they can be used in a discourse parser for automated discourse analysis in the line of Rhetorical Structure Theory (RST, Mann & Thompson 1988). Traditionally, cues for relational discourse analysis have been derived from lexical discourse markers and syntactic features of an input text. This strategy is well-established and has yielded good results for texts of a limited size such as newspaper articles. But discourse relations between larger segments of text such as the sections and paragraphs of a research article are frequently not signalled by cues that can be identified by shallow analyses of vocabulary and grammar. We thus suggest to additionally base discourse parsing on an analysis of the logical document structure

(the division of a text into textual objects such as titles, paragraphs, tables, and lists) and the generic document structure (parts of the text corresponding to text type-specific categories such as *introduction, method, results*, and *discussion* in the case of scientific articles). We aim to clarify what kind of cues and constraints for relational discourse analysis can be observed on these levels and how they can be derived from the corresponding linguistic and structural annotations of a text collection. To this end, we examine a corpus of German and English scientific articles in the fields of psychology and linguistics. Its documents are annotated on the three levels in question, namely their logical document structure, their generic document structure (text type structure) and a discourse structure according to RST (which is also the target structure of an RST discourse parser).

2. Representation of relational discourse structure

Rhetorical Structure Theory (RST, Mann & Thompson 1988; Marcu 2000) shares three basic assumptions with other linguistic discourse theories like *Segmented Discourse Representation Theory* (SDRT, Asher & Lascarides 2003) and the *Unified Linguistic Discourse Model* (ULDM, Polanyi et al. 2004a; Polanyi et al. 2004b): 1. Discourse structure can be modeled as a system of discourse coherence relations which hold between parts of text, i.e. elementary or complex discourse segments. 2. Complex discourse segments are structured hierarchically and can be represented either as a graph (SDRT) or as a tree (ULDM, RST). 3. Discourse coherence relations are either hypotactic (subordinating, mononuclear) or paratactic (coordinating, multinuclear).

RST can be considered a functional theory of text structure: "It describes the relations among text parts in functional terms." (Mann & Thompson 1988: 271). Relations between discourse segments are, amongst others, identified and described according to the effect the relational propositions have on the reader. The assignment of relations to pairs of discourse segments thus involves the recognition of the goals and beliefs of authors and readers about the meaning and function of these discourse segments. "An RST analysis always constitutes a plausible account of what the writer wanted to achieve with each part of the text. An RST analysis is thus a functional account of the text as a whole." (Mann & Thompson 1988: 258). The definitions of the RST relations reflect this functional perspective by describing the characteristics (i.e. constraints and effects) of the relations from the points of view of author and reader, as intentions and effects (see Table 1 for an example of a definition of an RST relation).

Table 1. Definition of the relation Concession (Mann & Thompson 1988, p.254f.)

relation name:	Concession
constraints on N:	W (the writer) has positive regard for the situation presented in N
constraints on S:	W is not claiming that the situation presented in S doesn't hold
constraints on the N + S combination:	W acknowledges a potential or apparent incompatibility between the situations presented in N and S; W regards the situations presented in N and S as compatible; recognizing that the compatibility between the situations presented in N and S increases R's positive regard for the situation presented in N
the effect:	R's (the reader's) positive regard for the situation presented in N is increased
locus of the effect:	N and S
discourse markers:	in S: obwohl, obschon, obgleich, obzwar, zwar [although, though ...] in N: dennoch, doch, trotzdem [however, anyhow, nevertheless …] in S: obwohl – in N: so in S: zwar – in N: jedoch

```
<hypo relname="concession" id="i217">
<s><dm lemma="zwar" pos="ADV">Zwar</dm> wurde in der Fremdsprachenerwerbsforschung im
   Zusammenhang mit der noticing-Hypothese die Rolle der auf den Input gerichteten Auf
   merksamkeit untersucht. </s>
<n>Die Funktion der lernerseitigen Aufmerksamkeit für den Output im L2-Erwerb blieb
   bisher <dm lemma="jedoch" pos="ADV">jedoch</dm>weitgehend unberücksichtigt. </n>
</hypo>
```

Listing 1. Discourse marker for Concession

When judgments about the intentions of the author play such a crucial role in discourse analysis according to RST, in which way can this theory be implemented in a computational approach to discourse parsing? In previous projects (Corston-Oliver 1998; Marcu 2000; Carlson & Marcu 2001; Reitter 2003), linguistic properties like syntactic or lexical features have been applied as *discourse markers* for the assignment of a discourse relation to sets of discourse segments. In this article, the term discourse marker is used to refer to a class of expressions (adverbs or conjunctions) as well as syntactic constructions which can be distinguished by their "function in discourse and the kind of meaning they encode" (Blakemore 2004: 221). Discourse markers are treated as signals the author uses to communicate his goals and beliefs to the reader, and which, more specifically, signal a specific discourse relation. Examples of discourse markers are connectives like "jedoch" (= "however"), which marks the nucleus of a Concession-relation (see Listing 1[1]), parallel syntactic constructions which may

1. The examples in this article are taken from:

Baßler, H. and H. Spiekermann (2001). Dialekt und Standardsprache im DaF-Unterricht. Wie Schüler urteilen – wie Lehrer urteilen. In: *Linguistik Online*, 9; Bühlmann, R. (2002).

induce a List-relation, and punctuation marks such as ':' (colon), which when occurring at the end of a segment may signal a Preparation-relation. As a consequence, an RST analysis does not, unlike an SDRT analysis, require a fully-fledged semantic representation of discourse segments.

Such a surface-oriented approach often works well for shortish text types, e.g. newspaper articles, which are characterised by a limited size and a relatively simple document and syntactic structure (Marcu 2000; Carlson & Marcu 2001; Reitter 2003). But when it comes to complex text types or longer texts with a deeply nested discourse structure, it is necessary to consider additional knowledge sources which can provide cues and constraints for the interpretation of higher levels of discourse structure, and discourse relations which are not indicated by lexical or syntactic discourse markers (such as Elaboration or Background). The term *cue* is used here to contrast with the term *discourse marker* to refer to more abstract signals for discourse relations like cues from the logical or generic document structure. We also distinguish cues from *constraints*: While cues can be used for bottom-up relational discourse analysis and in that respect behave like discourse markers that signal a specific discourse relation, constraints serve as top-down restrictions for discourse structures.

In our project (Lobin et al. 2010; Lüngen et al. 2006), we deal with a corpus of scientific articles[2] which exhibit a highly complex document structure and a relatively large average size (~ 8600 words per article). A complex document structure implies that they are characterised by a deeply nested hierarchical structure with several levels of embedded discourse segments where the distance between the level of elementary discourse segments (EDS) and the highest level of complex discourse segments (CDS) may be five or more embeddings. As a consequence, the majority of discourse segments are not elementary, but complex – this means that lexical and syntactic discourse markers can only be applied in a limited way. Apart from their complex document (and discourse) structure, our corpus of scientific articles is characterised by a high frequency of Elaboration relations – which are usually not indicated by lexical or syntactic discourse markers.

Ehefrau Vreni haucht ihm ins Ohr... Untersuchungen zur geschlechtergerechten Sprache und zur Darstellung von Frauen in Deutschschweizer Tageszeitungen. In: *Linguistik Online*, 11; Bärenfänger, O. and S. Beyer (2001). Zur Funktion der mündlichen L2-Produktion und zu den damit verbundene kognitiven Prozessen für den Erwerb der fremdsprachlichen Sprechfertigkeit. *Linguistik Online*, 8.

2. The whole corpus comprises 120 scientific articles in different languages (English and German), domains (psychology and linguistics) and sub-genres (experimental and review); this corpus is split in two subcorpora: PsyEngl (English psychology articles) and LingDeu (German linguistic articles). Part of the work described in this article (e.g. the canonical sequence of global text type structure categories) is based on the smaller LingDeu corpus of 47 linguistic articles.

Originally, RST provides a set of 26 rhetorical relations, which are either mono- or multinuclear (Mann & Thompson 1988). Apart from the distinction between mono- and multinuclear relations, relations can also be subdivided into two groups based on the intentions of the author and the effects on the reader: 1. subject matter relations: "those whose intended effect is that the reader recognises the relation in question"; 2. presentational relations: "those whose intended effect is to increase some inclination in the reader, such as the desire to act or the degree of positive regard for, belief in, or acceptance of the nucleus." (Mann & Thompson 1988: 257).

For our corpus of scientific articles, we adapted the relation set proposed by Mann and Taboada (2005) but extended and restructured it – by defining our own relation taxonomy (as was also done by Hovy & Maier 1995; and Carlson & Marcu 2001). The motivation for all modifications of the relation set is twofold: First, the set had to be adapted to the characteristics of the text type of the documents in our corpus. Second, the set should support our application scenario[3] and therefore distinguish between relations that are mainly induced by the logical document structure or generic document structure, and relations that are mainly induced by lexical, syntactic or morphological features.

A subset of 20 articles of our corpus with approximately 172.240 words was annotated according to rhetorical structure using the RSTTool developed by O'Donnell (2000). Subsequently, the representations produced were converted by a Perl script into our *RST-HP* format (Lüngen et al. 2006), where, unlike in O'Donnell's representation, the basic XML tree structure of a document is also used to represent an RST tree. An RST-HP extract is displayed in Listing 1.

Besides logical and generic document structure, another level which might be called *thematic structure* plays an important role in the instantiation of certain "subject matter" relations like Background, Elaboration, and its subtypes (Lüngen et al. 2006). In the present article, however, we will solely concentrate on constraints and cues that can be derived from the logical and generic document structure.

3. Cues and constraints from logical document structure

According to Power et al. (2003: 213), "document structure describes the organization of a document into graphical constituents like sections, paragraphs, sentences, bulleted lists, and figures" as well as elements like "quotation and emphasis". Such

3. The application scenario of our project is an online-system which supports novice learners (first or second year students) in selective and efficient reading of scientific articles, and which can furthermore be used as a learning environment where students may learn something about the structural and argumentative characteristics of the genre "scientific article". In order to personalise the system and to allow students to upload scientific articles, a discourse parser is being developed which implements the task of automatically adding discourse structure annotations.

constituents can be described according to their graphical or geometric properties – they are 2D-objects which cover parts of the document area (Lobin et al. 2010). The *physical layout* structure of a document is a manifestation of its *logical* document structure, as in the physical layout different structural functions of parts of a text can be identified such as list, paragraph, or heading. In our corpus, logical document structure is encoded according to DocBook markup (Walsh & Muellner, 1999).

At the level of the logical document structure, we can distinguish elementary and complex constituents. The latter are combinations of adjacent elementary or (smaller) complex constituents (parts of text). This combination follows compositional principles. A document can therefore be described as structured hierarchically: complex constituents are aggregations of elementary or complex ones, e.g. an article consists of sections, a section is divided into a set of subsections or paragraphs and perhaps lists or figures, and a paragraph may contain quotations or emphasised tokens. Elements of the logical document structure indicate specific discourse substructures and can thus be used as cues for the assignment of rhetorical relations. Document structural elements which serve as cues are, for example, *listitem, glossterm, glossdef, caption* and *title*. *Listitems* indicate a List_dm-other relation between all items of a bulleted list (as shown in Listing 3 and Figure 2), *glossterms* may induce the nucleus in an Elaboration-definition relation, *glossdef* the satellite (shown in Listing 2 and Figure 1), *captions* often have the status of a satellite in a Circumstance relation with a figure or table being the nucleus, and *titles* are the satellite in a Preparation-title relation, where the nucleus may be a section, a table or a figure.

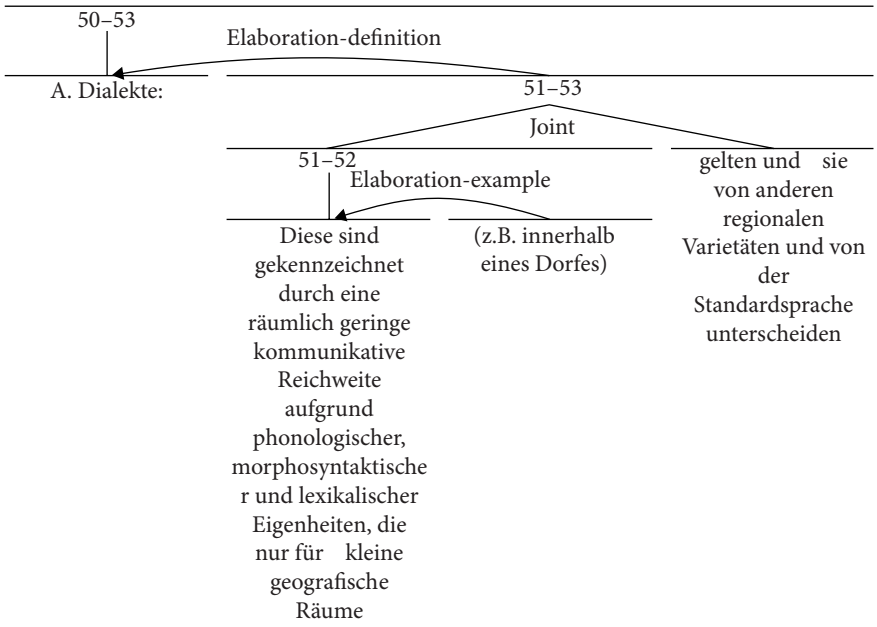


Figure 1. RST annotation for extract in Listing 2

```
<glosslist>
  <glossentry>
    <glossterm>A. Dialekte: </glossterm>
    <glossdef>
       <para>Diese sind gekennzeichnet durch
          eine räumlich geringe kommunikati
          ve Reichweite aufgrund phonologi
          scher, morphosyntaktischer und le
          xikalischer Eigenheiten, die nur
          für kleine geografische Räume
         (z.B. innerhalb eines Dorfes) gel
          ten und sie von anderen regionalen
          Varietäten und von der Standard
          sprache unterscheiden.
       </para>
    </glossdef>
  </glossentry>
</glosslist>
```

Listing 2. DocBook annotation (extract)

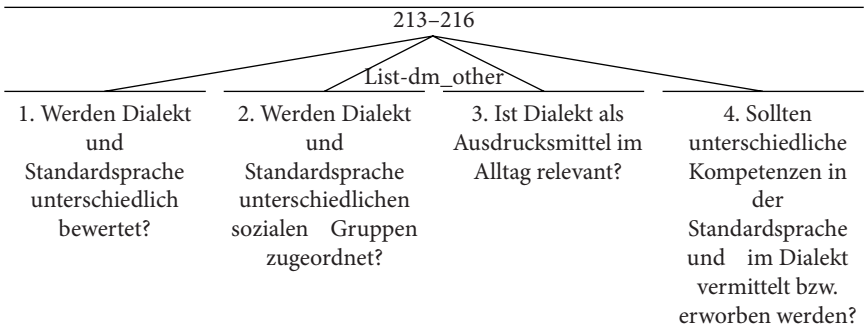


Figure 2. RST annotation for extract in Listing 3

```
<orderedlist>
    <listitem>
        <para>1. Werden Dialekt und
        Standardsprache unterschiedlich
        bewertet?</para>
    </listitem>
    <listitem>
        <para>2. Werden Dialekt und
        Standardsprache unterschiedlichen
        sozialen Gruppen zugeordnet?</para>
    </listitem>
    <listitem>...</listitem>
</orderedList>
```

Listing 3. DocBook annotation (extract)

Apart from these (and other) *cues*, the logical document structure also provides *constraints* for relational discourse analysis, insofar as the units of the logical document structure act as building blocks for discourse spans. Before we explain this statement in more detail, we shortly have to introduce our typology of discourse

segments. Because of the complexity and deep nesting of discourse segments in scientific articles, we do not only distinguish elementary (EDS), sentential (SDS), and complex discourse segments (CDS), but we additionally distinguish different types of CDS (see Figure 3 for a graphical illustration):

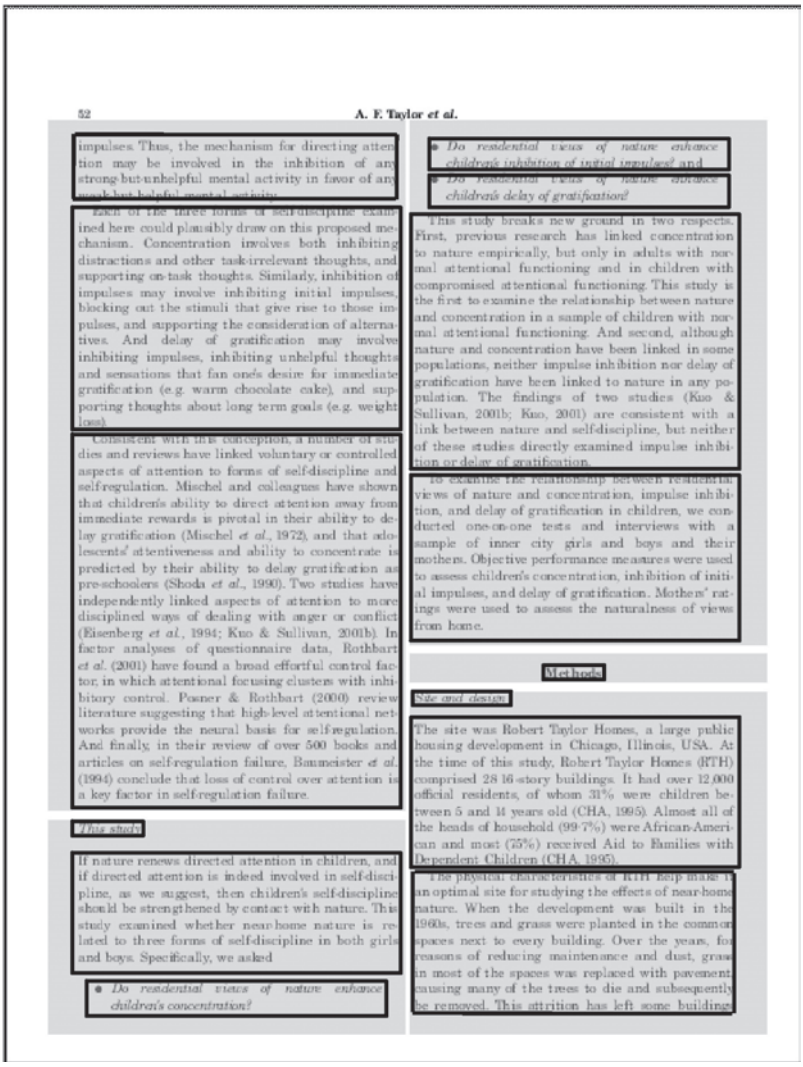


Figure 3. Areas of text that correspond to complex discourse segments of the types *block* (black frames) and *division* (grey rectangular areas)

– CDS type = "block": This segment type corresponds to paragraphs and structural elements that are on a par with paragraphs such as titles and captions, i.e. all element types from the logical structure that contain only text or text plus

inline elements. The name of the attribute "block" of this complex segment type is due to its correspondence to geometric objects that are two-dimensional blocks (rectangles) in the physical layout of a document. The segments of type "block" partition the document, i.e. every piece of text is part of exactly one CDS type = "block". In a discourse tree for a CDS type = "block", the block acts as an upper boundary and top node, and SDS (sentential discourse segments) act as the basic segments for the construction of discourse subtrees.

- CDS type = "division": Complex discourse segments of the type "division" correspond to the lowest section level. In terms of DocBook markup it comprises the smallest occurring sect1, sect2, sect3, sect4, or sect5 elements plus elements that are on a par with it, i.e. titles and paragraphs that are sisters of sect elements. The segments of type "division" also partition the whole document. In a discourse tree for a CDS type = "division", the division acts as an upper boundary, and the CDS type = "block" acts as a basic segment type for the construction of discourse subtrees.
- CDS type = "document": This segment type comprises all residual sect elements, i.e. those which are on a higher level than the ones described under CDS type = "division".

Thus, the CDS type = "document" level is special in that its segments most of the time do not partition the document, depending on the depth of embedded sections. On the other hand, the segments of the "document" can be identical to those of the type "division" in a document that contains only sections of the DocBook element type sect1. In a discourse tree for a complete document, the segments of type "division" are the basic elements, and all segments of the type "document" must correspond to exactly one subtree.

This differentiation of levels or granularities of discourse segments is comparable to that proposed by Marcu (2000), who distinguishes clause, sentence, paragraph, and section level, and LeThan et al. (2004), who describe sentence-level and text-level discourse segments. In our approach, units of the logical document structure (paragraph, section, article) are used to constrain the extent to which discourse segments can be relationally combined to pairs of discourse segments, i.e. they serve as boundaries for discourse segments. This means that, for example, a CDS type = "block" can only be related to another CDS type = "block", but not to a CDS type = "division".[4] By assuming that the rhetorical structure correlates with the logical document structure, or, as Marcu (2000: 109) says, "that sentences,

4. This constraint is of course an idealization to some degree. In the annotation process, for example, we noticed that sometimes the last sentence(s) of one paragraph serves as a preparation for the content of the following paragraph(s) so that it seems to be rhetorically closer to the following paragraphs.

paragraphs, and sections correspond to hierarchical spans in the rhetorical representation of the text", the amount of possible rhetorical interpretations can be reduced significantly.

In the following sections, we will examine the generic document structure of scientific articles and how it can function as a knowledge source in discourse parsing according to RST. First, a text type structure schema for scientific articles and the corpus annotations based on it will be presented. Next, we investigate what kind of cues and constraints can be derived from the generic structure and from statistic analyses of the annotations.

4. Cues and constraints from generic document structure

4.1 Interrelations between generic document structure and relational discourse structure

Apart from the *logical* document structure, a second type of document structure exists: the *generic* document structure, or genre-specific *text type structure* (TTS), or *superstructure* (Swales 1990; van Dijk 1980). It describes the global organisation of a document into genre-specific functional categories (or *zones*, after Teufel 1999) like, for example, *Problem, Method,* and *Results* (= categories of scientific articles). These categories represent functions of parts of a text as an instance of a specific text type, which are oriented towards the text as a whole. They can be organised hierarchically and therefore be formally described by a hierarchical schema (e.g. Kando 1999). The text type structure schema we developed for linguistic scientific articles is shown in Figure 4. The schema is based on the ones by Kando and by Teufel, but was, as a result from our corpus analyses, adapted to our corpus and to our application scenario. The global TTS categories for scientific articles are *Problem, Background, Evidence* (*Framework* and *Method*) and *Answers*. These categories can be further distinguished in *OthersWork, ResearchTopic, Theory, Data, Results* etc. In an earlier version, the schema consisted of 135 categories. When building the reduced schema with only 17 leaf categories that was used in this study, several of the original more fine-grained categories were combined to form broader categories. A suffix "_R" in the category label in Figure 4 indicates that the present category comprises several (residual) categories that were also found under the same super-category in the original schema.[5]

5. See Bayerl et al. (2003) for more information about the extended schema.

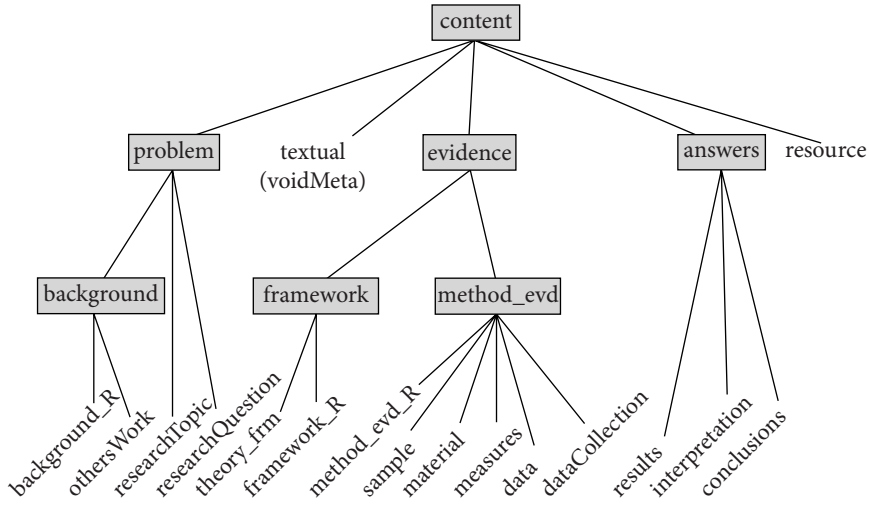


Figure 4. Text type structure (TTS) schema for scientific articles

For the annotation of our corpus each text was divided in TTS segments, often, but not necessarily always, consisting of a sentence. This (more or less) sentential segmentation corresponds to the segmentation realised by Kando (1999) and Teufel (1999). An example of a sentential TTS annotation is shown in Listing 4. Apart from the sentential TTS structure which we consider as the *local* text type structure, we additionally identify a macro TTS structure, or *global* text type structure. Each scientific article can be divided into functionally coherent macro sections which can be categorised according to the set of global TTS categories (Figure 4 – global categories are indicated by the grey boxes). Empirical analyses showed that the majority of scientific articles have a more or less canonical TTS structure which means that they are composed as a sequence of *Problem, Evidence, Answers*. Nevertheless, there are several articles which exhibit various deviations from this canonical structure. Not all global categories have to be present in scientific articles and the sequence of categories may vary, too. These articles can therefore be described as generic variations (see Section 4.3).

```
<segment id="s196" parent="g4" topic="results">In den Texten ist sehr oft
   nicht klar, ob ein Maskulinum nur auf Männer oder auch auf Frauen refe
   riert. </segment>
<segment id="s197" parent="g4" topic="interpretation"> Wichtige Fragen, die
   die LeserInnen an den Text haben, bleiben somit unbeantwortet. Die Politik
   wird durch den fast durchgehenden Gebrauch des generischen Maskulinums als
   "Männersache" dargestellt, Frauen werden, auch wenn sie vorhanden sind,
   selten sichtbar gemacht. Zudem wird auch mit geschlechtsspezifisch männli
   chen Wörtern wie Gründerväter der Gedanke an Männer evoziert. </segment>
```

Listing 4. TTS annotation (extract)

The role of the generic document structure for discourse analysis in the tradition of RST has lately been examined by Gruber and Muntigl (2005), and Taboada and Lavid (2003). Both approaches model the generic structure of a document as genres and stages (like *Orientation, Background, Account, Interpretation, Summary*) in the tradition of the Register and Genre Theory, i.e. as serially occurring functional stages, where each stage depends on a previous stage. Gruber and Muntigl empirically show that generic and rhetorical structure of students' academic writings coincide.[6] They found correlations between both genre dependent and independent stages, and RST relations. *Orientation*, for instance, typically occurred with Preparation, *Discussion* with Background and *Summary* with Summary (Gruber & Muntigl 2005: 102). These systematic relationships between generic and rhetorical structure are differentiated for different textual levels (and generic stages), i.e. for high level textual structures (stages) as well as low levels (substages).[7]

Likewise, Taboada and Lavid provided empirical evidence for correlations between generic stages and rhetorical (and thematic) patterns in scheduling dialogues, e.g. *Opening* correlated with Solutionhood, *Closing* with RST relations like Evaluation, Restatement, and Summary. Their intention was to use rhetorical relations as signals for a specific generic stage. Our approach works just the other way round. We intend to use the existing (so far manually assigned) generic document structure as a signal for a specific discourse structure. In our approach, three different ways of using the generic document structure as a cue or constraint for discourse interpretation are distinguished:

1. A TTS category which corresponds to an RST relation can be used as an explicit cue for a specific RST relation: As pointed out above, a text type structure category is a functional relation between a part of a text and the text as a whole, while an RST relation establishes a functional relation between two or more parts of a text (discourse segments). The category names for both types of functional relations, however, partly overlap, e.g. *Background* – Background, *Problem* – Problem-solution, *Evidence* – Evidence, Results – Result, *Interpretation* – Interpretation, and maybe also Summary – *Conclusions*. We suspect that TTS categories are also often functions between parts of a text rather than

6. Their corpus consists of 19 student academic term papers (lengths ranging between 1865 and 7271 words). For the annotation, 35 RST relations were used and 46 genre stage categories.

7. Gruber and Muntigl's relational discourse analysis based on RST was restricted to the level of subchapters; clauses were not annotated.

between a part and the text as a whole. The *Answers* section of a scientific article, for example, contains an answer to what is described in the *Problem* part rather than an answer to what is described in the text as a whole. Hence, it seems reasonable to identify certain TTS categories with equivalent RST relations, e.g. an *Interpretation* TTS constituent should be an RST satellite in an Interpretation relation. Empirical evidence for this hypothesis is provided by the findings of a descriptive analysis of our corpus.[8] The relation Interpretation is found 17 times as frequently in TTS segments which are of type *Interpretation* than with all other TTS categories (on average), Background occurs 9 times as frequently with *Background*, and Summary 9 times as frequently with *Conclusions* (see Section 4.3).

2. Generally, a TTS category (assigned to a TTS segment) which frequently appears with RST relation A and never with relation B induces relation A with a higher probability than relation B – the TTS category can therefore be used in a statistic constraint: The quantitative analysis of our corpus – similar to the empirical research done by Gruber and Muntigl (2005) – showed high deviations from the average distribution of relations and TTS categories. Some TTS categories correlate significantly with one or two specific RST relations. The analyses and its findings will be described in greater detail in Section 4.3.
3. At the highest level of discourse structure (CDS type = "document"), the global categories of the text type structure schema (*Problem*, *Background*, *Evidence*, *Framework*, *Method*, *Answers*) should be determined automatically for all CDS type = "division", so that the relations between these categories can be inferred using a relational schema (see Figure 5). This approach is based on the fact that in most cases, scientific articles are organised along a specific sequence of global generic categories. For a detailed description of the procedure of instantiating global TTS categories and relations between them see Section 4.2.

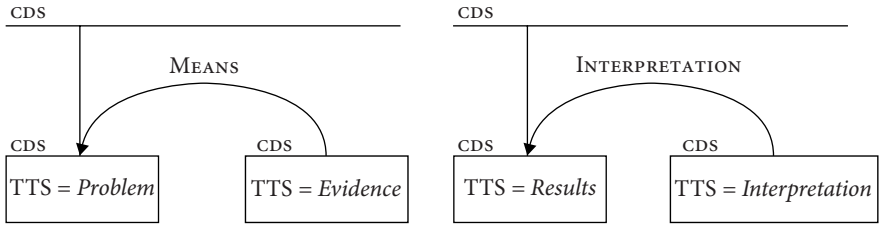


Figure 5. Rhetorical relations induced by TTS annotations of adjacent complex discourse segments

8. 2 x 10 texts of our corpus were analysed: For the 10 German linguistic articles, RST annotations were done on the level of paragraphs (CDS type = "block") as minimal units, for the 10 English psychology articles RST annotations were done with clauses (EDS) as minimal units.

4.2 Canonical sequence of global text type structure categories

The offline instantiation of global TTS categories for all top-level divisions (DocBook: sect1-n) was based on an analysis of the size (measured in the number of tokens contained)[9] of all sentential TTS categories of one section. In the analysis, it was calculated which of the global categories comprised the largest local categories, i.e. the parent category (in the tree-structured text type structure schema shown in Figure 4) of the majority of local TTS categories in the current section was looked up. Each section was then labelled with the global category found such that the whole article is annotated as a sequence of macro section segments with a TTS category assigned (see Listing 5).

```
<segment id = "i4" topic = "problem" strtype = "sect1">0 Einleitung ... </segment>
<segment id = "i5" topic = "framework" strtype = "sect1">1 Positionierung des Pro
   jektes im Forschungskontext ... </segment>
<segment id = "i6" topic = "researchTopic" strtype = "sect1">2 Erkenntnisinteressen
   und Ziele ... </segment>
```

Listing 5. Global TTS annotation (extract)

To determine whether the canonical sequence of TTS categories expressed in the text type structural schema in Figure 3 holds for the scientific articles in our corpus, we calculated the sequences of global TTS categories of each article in the corpus. We found one basic TTS sequence (Type A), which may be described as the generic prototype, and two variations of this prototype (Types B and C):

– **Type A**: This type is regarded as the prototypical one for scientific articles. Articles of this kind exhibit a sequence of all global categories *Problem* – *Evidence* (which may be split into *Framework* and *Method*) – *Answers* (which may be split into *Results* and *Interpretation*), (26 articles out of 47). The initial section of an article is always *Problem* and the final one is always *Answers*. By contrast, the central part of scientific articles is much less restricted than the beginning and the end. Disruptions of the canonical order occur frequently, for example through sections annotated as *Others Work, Background* or *Framework* (7 articles of the 26). Especially the latter three categories seem to be sequentially more variable than other categories like *Conclusions* or

9. We would like to thank Mario Klapper who did most of the manual XML annotations and numeric analyses for the work presented in this section of the paper.

Method. In one case, an article had an initial section labelled as *Answers*, because it described the findings of the study presented. It is not unusual for sentential *Answers* segments to occur in the first section of an article, but it is atypical that these segments constitute the major part of the first section. The case is therefore treated as an (atypical) variant of A.

- **Type B**: Under this type we subsume articles which exhibit the basic sequence of global categories but with one of *Problem* or *Evidence* or *Answers* missing (12 articles). Type B is the most common alternative to Type A. In most cases (8 articles), *Problem* does not occur as an initial section of its own. Instead, TTS segments with local subcategories of *Problem* are interspersed in various sections, but in none of them do they constitute the majority of sentential TTS segments.
- **Type C**: Articles whose sections are all annotated with the global category *Problem:Background* (3 articles), or with *Evidence:Framework* (1), or articles which are annotated with an arbitrary sequence of the *Problem* and *Framework* categories (3). These structures are regarded as atypical alternatives to the canonical text type structure and are mainly found in articles about theoretical work or language politics.
- **Deviations**: There is a small subset of articles (2 articles) which show various deviations from the structures introduced in Types A–C. We treat these articles as special cases.

The high number of articles in Types A and B confirms the relevance of the text type structure schema in Figure 4 and the canonical order of categories expressed in it. As a consequence, when complex discourse segments of the highest type (CDS type = "division") are annotated with TTS categories, relations between them can be instantiated with a certain confidence. The corresponding configurations of TTS segments thus serve as discourse structural cues as illustrated in Figure 5.

4.3 Correlations between text type structure categories and rhetorical relations

The subcorpus used in this study comprises two parts: The first consists of 10 English psychology articles with 8597 words on average. In this part (henceforth: PsyEngl), the minimal segments are elementary discourse segments (EDS), so RST relations are annotated for the EDS level and all higher levels. The second part of our corpus contains 10 German linguistic articles with 8627 words on average – this subcorpus (henceforth: LingDeu) is annotated with rhetorical relations

between complex discourse segments of the type "block" and higher. On both subcorpora we examined correlations between RST relations and TTS categories. For this task, we employed the *Sekimo Tools* for the analysis of multiple XML annotations of one textual base (Witt et al. 2005). We made use of two features of the Sekimo Tools:

1. *Markup unification*: For each article, we produced a unified XML document containing both its logical document structure annotation and its TTS annotation of elementary TTS segments. By means of an XSLT style sheet, we then automatically assigned TTS categories to the *TTS block segments* of an article based on the existing TTS annotations of elementary TTS segments. The TTS category which in terms of the number of word tokens made up the largest part of a block segment was selected as the TTS category for that segment as a whole. Subsequently, adjacent block segments with identical TTS categories were joined into one TTS segment.
2. *Relation checking*: Using the Sekimo Tools, inferences about the relationship between elements on an annotation layer A and another annotation layer B of one textual base can be drawn. Possible relations between elements on different layers are *inclusion, overlapping, adjacency, independence* and others (determined in terms of the shared or unshared PCDATA element contents). Thus, taking both the newly obtained TTS block segment annotation and the RST annotation of each document, we listed the rhetorical relations between discourse segments *included* in the TTS block segment and related them to the TTS category of that segment.[10]

As a result, we obtained a contingency matrix with the relations arranged in lines and the TTS categories in columns. The matrix shows the type and the number of relations for each TTS category. Table 2 shows the frequencies of various TTS categories and RST relations, and the number of TTS segments and included RST segments for both. The percentages of the most frequently occurring TTS categories in terms of their share in the total number of TTS segments (121 in the PsyEngl and 361 in the LingDeu subcorpus) are *Framework* (20%), *Results* (17%) and *Measures* (11%) in PsyEngl, and *Data* (25%), *Results* (25%) and *Framework* (12%) in LingDeu.

10. The RST annotation also contains numerous relation instances *not* included in a TTS segment, i.e. relations between segments of larger types than CDS type = "block".

Table 2. Number and frequency of TTS categories and RST relations in the corpora

	#TTS category types	#TTS segments[11]	Most frequent TTS segment categories	#RST relation types	#RST Segments included in TTSs	Most frequent TTS categories for RST segments (as p(TTS))	Most frequent RST relations included in TTS (as p(Rel))
PsyEngl corpus	17	121	*Framework*: 20%	36	801	p(*Framework*) = 25%	p(Elaboration) = 35%
			Results: 17%			p(*Measures*) = 21%	p(List) = 9%
			Measures: 11%			p(*DataCollection*) = 11%	p(Circumstance) = 8%
LingDeu corpus	17	361	*Data*: 25%	17	297	p(*Results*) = 37%	p(List)=33%
			Results: 25%			p(*Data*) = 14%	p(Elaboration) = 23%
			Framework: 12%			p(*Framework*) = 10%	p(Preparation) = 23%

For our following analyses, we took the total number of RST segments included in a TTS segment and annotated with an RST relation (801 for PsyEngl and 297 for LingDeu) to calculate the percentages that express the estimated a priori probabilities *p(TTS)* (the probability of an RST segment to be included in a specific TTS) and *p(Rel)* (the probability of an RST Segment to be annotated with relation Rel). The highest percentages for different TTS are p(*Framework)* = (202/801)*100 = 25%, p(*Measures)* = 21%, and p(*DataCollection*) = 11% in PsyEngl, and p(*Results)* = 37%, p(*Data)* = 14% and p(*Framework)* = 10%. in LingDeu. The percentages or a priori probabilities p(Rel) for the most frequent RST relations are p(Elaboration) = 35%, p(List) = 9%, p(Joint) = 8% and p(Circumstance) = 8% in PsyEngl, and p(List) = 33%, p(Elaboration) = 23%, and p(Preparation) = 23% in LingDeu.

In a subsequent step, we examined the distribution of relations in each TTS category and listed them as percentage values which correspond to the conditional probability for the occurrence of an RST Segment with relation *Rel* given a Text Type Structure category *TTS*, i.e. *p(Rel|TTS)*. For example, p(Elaboration|*Framework*) = 38%, p(Joint|*Framework*) = 10%, p(Condition|*Framework*) = 8%, and p(List|*Framework*) = 5% in the PsyEngl corpus. For each relation type, we then calculated the average percentage of its

11. Due to their small size, some TTS segments do not contain any RST segments.

frequency (afp) over all TTS categories. The average frequencies of RST relations over all TTS categories are:[12]

- For PsyEngl: afp(Elaboration) = 30.2%, afp(List) = 13.8%, afp(Circumstance) = 9.8%, afp(Joint) = 8.2%, afp(Preparation) = 7.2%
- For LingDeu: afp(Preparation) = 28.8%, afp(List) = 27.9%, afp(Elaboration) = 18.1%, afp(Summary) = 9.1% , afp(Evidence) = 5.2%

To find out which RST relations are more prominent in TTS category A than in TTS category B, we compared the average percentage with the actual percentage of each relation Rel for a specific TTS category by calculating the deviation factor D*(Rel, TTS) : = p(Rel|TTS)/afp(Rel)*. For example, D(Condition,*Framework*) = 8/1.44 = 5.6, i.e. Condition can be found 5.6 times more often in *Framework* than (on average) in all other TTS categories. The results of all these calculations are given in Figure 6. The diagram shows the different distributions of RST relations. The peaks in the graphs indicate that a TTS category can be clearly distinguished by a different distribution of RST relations. All TTS categories have special characteristics with respect to the occurrences and frequencies of RST relations – some relations are found up to 17 times more often in a specific TTS category than in any other one.

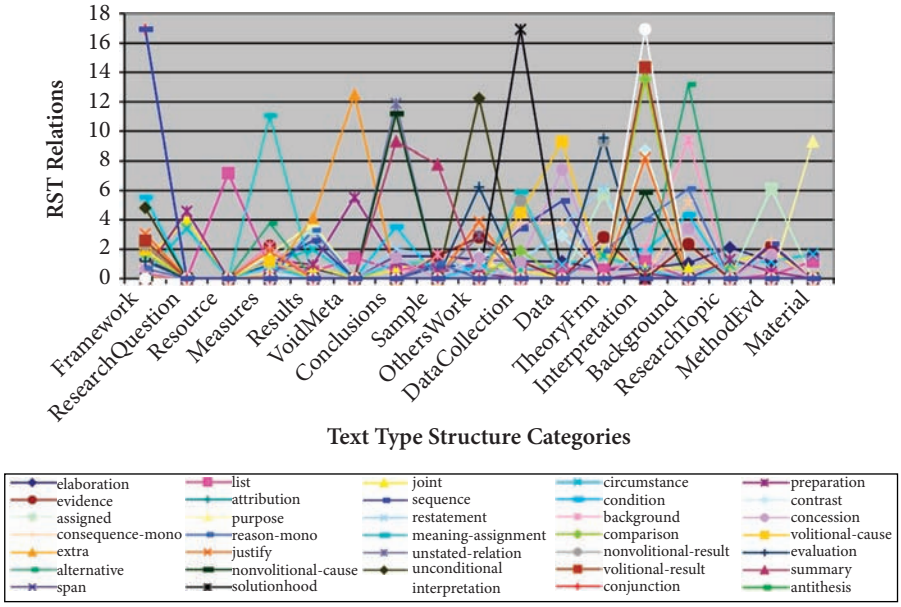


Figure 6. Distribution of RST relations over the different TTS categories for the PsyEngl corpus

12. These percentages do *not* refer to the number of RST relations in the corpus, but to the average distribution of relations in each TTS category.

Similarly, we calculated the deviation of the expected frequency of a TTS category at a specific RST segment from its actual frequency by calculating the factor E*(TTS, Rel) : = p(TTS|Rel)/p(TTS)*. For example, the percentage of ELABORATION segments under *Framework* w.r.t. the total number of ELABORATION segments is p(*Framework*|ELABORATION) = (76/282)*100 = 27%. The factor E(*Framework*, ELABORATION) = 27/25 = 1.1 describes the deviation between the expected number and the actual number of segments under *Framework* that are annotated with ELABORATION.

Some of the factors we obtained were extremely high. Often this was due to only a small number of included relation instances of a type in the corpus. For this reason, we ignored those relations which have less than 10 included instances in both subcorpora. Due to the different number of RST instances in the two corpora, the number of different RST relations with 10 or more included instances varies across the corpora. In the LingDeu corpus, only 10 RST relations have more than 10 included instances, whereas 19 relations have more than 10 included instances in the PsyEngl corpus.

It is remarkable that the findings for the two corpora are only partly overlapping. The reason for the differences could be either the different sizes of the minimal discourse segments (EDS for PsyEngl, CDS type = "block" for LingDeu), or the domain, or even a language-specific style of discourse organisation. One or all of these factors seem to influence the prominence of TTS categories and RST relations, e.g. English psychology articles contain many more segments of the TTS category *Measures* than German linguistic articles, whereas e.g. the RST relation PREPARATION is much more common in German linguistic articles. Therefore, it may be problematic to transfer the findings to scientific articles from other languages, domains and/or of different segment granularity. However, some of the correlations of TTS categories and RST relations are similar for both corpora and have therefore high empirical evidence. These are *ResearchTopic* – ELABORATION, *Sample* – SUMMARY, *Data* – CONTRAST, *Background* – CONSEQUENCE-MONO, *Background* – CONCESSION, and *Background* – BACKGROUND, cf. Table 3.

Apart from the correlations that both corpora exhibit, correlations that hold only in one corpus can be found. The clearest correlations of RST relations and TTS categories are those where both factors (Factor E and Factor D) are higher than 5.0. For the LingDeu corpus, these are *Measures* – SUMMARY, *Conclusions* – CONSEQUENCE-MONO, *OthersWork* – CONCESSION, and *OthersWork* – BACKGROUND. For PsyEngl, the most prominent correlations are *Conclusions* – SUMMARY, *Data* – CONCESSION, *Interpretation* – CONTRAST, *MethodEvd* – ASSIGNED, *TheoryFrm* – ASSIGNED, and *TheoryFrm* – RESTATEMENT.

Provided that the assignment of TTS categories to CDS segments of a scientific article can be done automatically (as has so far been tested and reported in Langer et al. 2004), the factors D calculated from our corpus can be employed as statistic

Table 3. Correlations between TTS categories and RST relations

LINGDEU/ PSYENGL	FW	MES	RES	CON	SMP	OW	DC	DT	INT	BCK	RT	ME	TF
Elaboration	2.2/…										**5.5/2.1**		
											4.4/…		
List								2.2/…					
Evidence					4.8/…	…/2.8							…/2.8
					5.3/…	…/4.9							
Summary		**11.0/…**		**…/9.3**	**2.8/7.7**		…/…						
		16.5/…		**…/10.3**	**4.1/8.5**		2.8/…						
Contrast							7.5/…	**7.6/2.9**	**…/8.7**				
							4.9/…	**5.1/2.1**	**…/6.4**				
Consequence-mono				**6.7/…**			…/3.7	…/3.8		**9.2/5.2**		…/2.5	
				5.5/…			…/3.2	…/3.3		**7.6/4.6**		…/2.2	
Reason-mono			…/3.3						…/3.9	…/6.1			…/…
			…/4.9						…/5.9	…/9.1			…/2.8
Concession						**8.0/…**		**…/7.4**		**8.0/3.4**			
						7.6/…		**…/9.1**		**7.6/4.2**			

Background			4.5/…		**6.2/…**				**6.2/9.4**			
			5.5/…		**7.6/…**				**7.6/8.3**			
Assigned			…/3.4								**…/6.1**	**…/5.8**
			…/3.8								**…/6.9**	**…/6.6**
Meaning-assignment		…/11.1				…/5.9						
		…/3.7										
Sequence			…/2.6			…/3.5	…/5.3					
			…/2.2			…/2.9	…/4.6					
Attribution					…/2.8	…/5.0				…/3.4		
						…/3.4				…/2.3		
Condition	…/5.5			…/3.6				…/4.3				
	…/2.8											
Restatement							…/3.6					**…/6.1**
							…/3.0					**…/5.1**
Legend:	FW = *Framework*		RES = *Results*	SMP = *Sample*		DC = *Data-Collection*		INT = *Interpretation*		ME = *Method-Evd*		
	MES = *Measures*		CON = *Conclusions*	OW = *Others-Work*		DT = *Data*		BCK = *Background*		TF = *Theory-Frm*		

constraints for the assignment or disambiguation of RST relations to segments pairs included in the CDS. Conversely, the factors E could be used if TTS assignment were the main goal of analysis using RST annotation as an auxiliary analysis. In our discourse parser, the former strategy will be pursued.

5. Conclusion and outlook

The aim of the study presented was to examine what kind of cues and constraints for discourse interpretation can be derived from the logical and generic document structure of complex texts such as scientific journal articles. Consequently, we performed several analyses on a corpus of scientific articles that is annotated on different XML annotation layers: The XML annotation of the logical document structure is realised by using an (extended) subset of the DocBook DTD (Walsh & Muellner 1999). The generic document structure (text type structure) is encoded using an XML schema based on the text type structure schema for scientific articles shown in Figure 4. Moreover, XML annotations of RST analyses of several articles were provided. So far, the texts of the corpus are annotated manually and semi-automatically.[13] The XML-based multi-layer annotation approach (Witt et al. 2005) was used to examine dependencies between XML elements on different annotation layers.

In Section 3, we argued that logical structure elements like *title, listitem, glossterm* can serve as cues in automated discourse analysis just like traditionally lexical discourse markers such as conjunctions and sentential adverbs. On the other hand, we introduced the discourse segment types EDS, SDS, and CDS (with its subtypes) and pointed out how they can be used as constraints to narrow down the textual domains inside which to identify relation spans. With regard to generic document structure, in Section 4.2 we showed that in our corpus a canonical sequence of text type structure categories occurring in the majority of articles can be established. Moreover, most deviations from this sequence could be grouped into two types. With a certain confidence, such sequences or partial sequences can be used as cues to assign discourse relations along relational schemas as presented in Figure 5. Finally, in 4.3, we demonstrated in a corpus analysis how and which

13. The DocBook annotation was produced semi-automatically, the annotation of the text type structure manually. A first approach to the automatic assignment of TTS categories to text segments in our corpus yielded mixed results: The recall and precision figures for some TTS categories were good, while recall for other categories was poor (Langer et al. 2004).

TTS categories assigned to complex discourse segments of type "block" statistically constrain the occurrence of rhetorical relation types.

In the future, we will work on integrating the cues and constraints described in this study into a discourse parser that takes several XML annotation layers of the same text as its input and provides a new XML annotation layer containing the discourse analysis as its output (Lüngen, et al. 2006). So far, the parser uses lexical discourse markers and several grammatical features as cues for relation assignment. The parser consists of cascaded iterative applications of a bottom-up chart parser and is realised in Prolog, and the discourse cues are encoded in the form of its *reduce rules*. These are derived from a discourse marker lexicon and also make reference to a syntax and morphology XML annotation layer which is generated using the commercial *Machinese Syntax* tagger software from Connexor Oy. Using this architecture, further reduce rules that make reference to the logical and the generic structure annotation layers of a document will be generated from the representations of the results of this study and integrated into the parser's rule files. Evaluations as to the contribution of this type of rules to the overall results will be provided.

References

Asher, Nicholas & Alex Lascarides. 2003. *Logics of Conversation*. Cambridge, UK: Cambridge University Press.

Bayerl, Petra Saska. & Daniela Goecke and Harald Lüngen & Andreas Witt. 2003. "Methods for the semantic analysis of document markup." In *Proceedings of the ACM-Symposium on Document Engineering (DocEng 2003)*, 161–170. Grenoble, France: INRIA.

Blakemore, Diane. 2004. "Discourse Markers." In *The Handbook of Pragmatics*, Laurence R. Horn & Gregory Ward (eds.), 221–240. Oxford: Blackwell.

Carlson, Lynn & Marcu, Daniel. 2001. *Discourse tagging reference manual.* [Technical Report ISI-TR-545]. Marina del Rey, CA: Information Science Institute.

Corston-Oliver, Simon. 1998. *Computing of Representations of the Structure of Written Discourse* [Ph.D. thesis]. Santa Barbara, CA: University of California.

Gruber, Helmut, Muntigl, Peter. 2005. "Generic and Rhetorical Structures of Texts: Two Sides of the Same Coin?" *Folia Linguistica* XXXIX (1–2):75–114. [Special Issue: Approaches to Genre].

Kando, Noriko. 1999. "Text structure analysis as a tool to make retrieved documents usable." In *Proceedings of the 4th International Workshop on Information Retrieval with Asian Languages*, 26–135. Taipei, Taiwan.

Holler, Anke. 2003. *Spezifikation für ein Annotationsschema für Koreferenzphänomene im Hinblick auf Hypertextualisierungsstrategien.* [http://www.hytex.uni-dortmund.de/hytex/publikationen.html#Dokus; retrieved 2009-08-06].

Langer, Hagen, Lüngen, Harald & Bayerl, Petra S. 2004. "Towards automatic annotation of text type structure: Experiments using an XML-annotated corpus and automatic text classification

methods." In *Proceedings of the workshop on XML-based richly annotated corpora (XBRAC) at the LREC 2004*, 8–14. Lissabon.

LeThanh, Huong, Abeysinghe, Geetha & Christian Huyck. 2004. "Generating Discourse Structures for Written Texts." In *Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004)*, 329–335. Geneva, Switzerland.

Lobin, Henning, Bärenfänger, Maja, Hilbert, Mirco, Lüngen, Harald & Puskas, Csilla. 2010. "Discourse relations and document structure." In *Linguistic modeling of information and Markup Languages, Contributions to language technology,* Dieter Metzing and Andreas Witt (eds.) [Series Text, Speech and Language Technology]. Dordrecht: Springer.

Lüngen, Harald, Bärenfänger, Maja, Hilbert, Mirco, Lobin, Henning & Puskas, Csilla. 2006. "Text parsing of a complex genre." In *Proceedings of the Conference on Electronic Publishing (ELPUB),* 247–256. Bansko, Bulgaria.

Mann, William. C & Thompson, Sandra A. 1988. "Rhetorical Structure Theory: Toward a functional theory of text organisation." *Text* 8(3): 243–281.

Mann, William C & Taboada, Maite. 2005. *Rhetorical Structure Theory. Relation Definitions*. [http://www.sfu.ca/rst/01intro/definitions.html; retrieved 2009-08-06].

Marcu, Daniel. 2000. *The theory and practice of discourse parsing and summarization*. Cambridge, MA: MIT Press.

O'Donnell, Michael. 2000. "RSTTool 2.4. A markup tool for Rhetorical Structure Theory." In *Proceedings of the International Natural Language Generation Conference (INLG'2000),* 253–256. Mitzpe Ramon, Israel.

Polanyi, Livia, Culy, Christopher, van den Berg, Martin H, Thione, Gian Lorenzo & Ahn, David. 2004a. "A rule based approach to discourse parsing." In *Proceedings of the 5th Workshop in Discourse and Dialogue*, 108–117. Cambridge, MA.

Polanyi, Livia, Culy, Christopher, van den Berg, Martin H, Thione, Gian Lorenzo & Ahn, David. 2004b. "Sentential structure and discourse parsing." In *Proceedings of the ACL 2004 Workshop on Discourse Annotation*, 49–56. Barcelona.

Power, Richard, Scott, Donia, & Bouayad-Agha, Nadjet. 2003. "Document structure." *Computational Linguistics,* 29(2): 211–260.

Reitter, David. 2003. "Simple signals for complex rhetorics: On rhetorical analysis with rich-feature support vector models." In *Sprachtechnologie für die multilinguale Kommunikation. Textproduktion, Recherche, Übersetzung, Lokalisierung. Beiträge der GLDV-Frühjahrstagung 2003*, Seewald-Heeg, Uta (ed.), 38–52 [Volume 18 of LDV-Forum].

Swales, John M. 1990. *Genre Analysis. English in academic and research settings.* Cambridge, UK: Cambridge University Press.

Taboada, Maite & Lavid, Julia. 2003. "Rhetorical and thematic patterns in scheduling dialogues." *Functions of Language* 10(2): 147–148.

Teufel, Simone. 1999. *Argumentative Zoning: Information Extraction from Scientific Text* [Ph.D. thesis]. University of Edinburgh.

van Dijk, Teun. A. 1980. *Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Walsh, Norman & Muellner, Leonard. 1999. *DocBook: The Definitive Guide,* Sebastopol, CA: O'Reilly.

Witt, Andreas, Lüngen, Harald, Goecke, Daniela and Sasaki, Felix. 2005. "Unification of XML documents with concurrent markup." *Literary and Linguistic Computing*, 20(1): 103–116.

Obligatory presupposition in discourse

Pascal Amsili[1] & Claire Beyssade[2]*
[1]Université Paris Diderot & Laboratoire de Linguistique Formelle, CNRS/[2]Institut Jean Nicod, CNRS Paris

Some presupposition triggers, like *too*, seem to be obligatory in discourses where the presupposition they induce is explicitely expressed. We show that this phenomenon concerns a larger class than is usually acknowledged, and suggest that this class corresponds to the class of presupposition triggers that have no asserted content. We then propose a pragmatic explanation relying on the neo-gricean notion of antipresupposition. We also show that the phenomenon has a complex interaction with discourse relations.

1. Introduction

The starting point of this work is a number of situations where some presupposition triggers seem obligatory in discourse. Here are two examples.

(1) a. Jean est allé il y a deux ans au Canada. Il n'ira plus là-bas.
John went to Canada two years ago. He won't go there anymore.

b. #Jean est allé il y a deux ans au Canada. Il n'ira pas là-bas.
John went to Canada two years ago. He won't go there.

c. Léa a fait une bêtise. Elle ne la refera pas.
Lea did a silly thing. She won't re-do it.

d. #Léa a fait une bêtise. Elle ne la fera pas.
Lea did a silly thing. She won't do it.

*This article benefitted from comments from the audience at CiD'06 (Maynooth), and particularly from discussion with Henk Zeevat. We are also very grateful to an anonymous reviewer who provided very detailed and constructive comments. Earlier versions of this work were presented at Toulouse (PICS France-Catalunya, June 06), Bordeaux (Signes, Oct 07), Carry-le-Rouet (ANR Prélude, June 07), Boston (Syntax-Semantics Reading Group, MIT, April 08) and Paris (Ling Lunch Paris Diderot, May 2009). We also wish to thank Grégoire Winterstein and Benjamin Spector for numerous discussions, and André Bittar for his help in preparing the final version of this paper. All errors remain our own.

In (1a-b), the presupposition trigger *ne... plus* (not anymore) is clearly preferred over the simple negation *ne... pas* (not).[1] However, both have the same asserted content, and the presuppositional content conveyed by *ne... plus* does not add anything in this particular context, because (1a) is a clear instance of presupposition binding (van der Sandt 1992; Kamp 2001): what *ne... plus* could add with respect to *ne... pas* (*i.e.* John went to Canada) is part of the common ground since it was already asserted in the first clause. In other words, we are here in a situation where the speaker seems to be "forced to presuppose", forced to use a presupposition trigger, even if this trigger doesn't bring any new information in the context.

To put it differently, we could say that in such cases, a form of informational redundancy seems obligatory, which is unexpected, redundancy being usually banned when, say, the same content is asserted twice, or even when already presupposed material is asserted. Here the redundancy has to be achieved by means of a presupposition trigger.

Our aim in this paper is first to show that the phenomenon, which has already been described in the literature for several particles, is more general than is usually acknowledged (Section 2), then to propose a pragmatic explanation for it (Section 3), and finally to take into account the interaction of the phenomenon with discourse in general (Section 4).

2. Data

We first survey in this section previous accounts of similar obligatoriness (§ 2.1), before trying to define the relevant class of presupposition triggers (§ 2.3), after having said a few more words on the importance of presupposition (§ 2.2).

2.1 Background: Obligatoriness of too and other additives

2.1.1 *Kaplan*

Although it was not presented exactly as we have just done, this phenomenon was first observed quite a long time ago, with respect to the "obligatoriness of *too*". According to (Kaplan 1984), this observation traces back to (Green 1968). The relevant examples include the contrast in (2).

1. It turns out that in French, *pas* and *plus* are really interchangeable and cannot occur together in such contexts. So we have a strong suggestion that they form an alternative, that the speaker has to choose between the two, whereas in English, for instance, the choice would be between adding *anymore* or not. But we believe that what we are dealing in this paper does not depend on this idiosyncratic property of French.

(2) a. Jo had fish and Mo did too.
b. *Jo had fish and Mo did. (Kaplan 1984, p. 510)

The simplest cases may suggest that we deal with syntactic constraints. However, semantics clearly plays a role, as can be seen with the next pair of examples. Here, *too* is not strictly obligatory, and the sentence (3b) is syntactically and semantically well-formed. However, it is pragmatically deviant, since it is suggesting that beeing seventeen is not beeing old enough to have a driver's licence.

(3) a. Barb is seventeen, and WENDY is old enough to have a driver's license, too.
b. #Barb is seventeen, and WENDY is old enough to have a driver's license.
(Green 1968)

Kaplan's proposal, in a nutshell, derives the obligatoriness of *too* from its discourse function, which is to "emphasize the similarity between members of a pair of contrasting items" (p. 516). This proposal relies crucially on the presence of a contrast, and applies only to examples like (2) where a conjunction (with *and* or *but*) is involved.

2.1.2 *Krifka*

In a paper about stressed additive particles, (Krifka 1999) makes several comments about the obligatoriness of *too*. It should be noted that his paper is concerned mainly with German focus-sensitive particles, in the particular case where they occur after the focus, as in (4).

(4) Peter invited Pia for dinner, tòo.

In such configurations, according to Krifka, the additive particle is always stressed (bearing a focus stress, noted with a grave accent), and it associates with a contrastive topic, itself stressed with a topic accent, noted with an acute accent.

His analysis of the reason why *too* is obligatory in such cases relies crucially on two facts:

1. the distinction between two types of accent, the focus accent, and the contrastive topic accent (following Büring's work (Büring 1998) and the classical distinction from (Jackendoff 1972) between A and B accents in English),
2. the existence of an implicature, derived from a distinctiveness constraint.

Let us recall that the placement of the focus accent is determined by the so-called discourse coherence constraint, which stipulates that the focus accent falls on the constituent which provides a congruent answer to the question (direct and exhaustive) as in (5).

(5) a. A: What did Peter eat?
b. B: Peter ate pàsta.
c. B′: *Péter ate pasta.

When the answer is partial, there is an additional accent, the topic accent, and it is obligatory:

(6) a. A: What did Peter and Pia eat?
b. B: *Peter ate pàsta.
c. B′: Péter ate pàsta.

Büring has shown that answers in which there is a topic accent are answers which leave open a number of questions. So for instance, in (6), the question of what Pia ate is left open. According to Büring, such uses of the topic accent are subject to a constraint called condition of disputability. Krifka claims that another constraint comes with contrastive answers, what he calls the distinctiveness constraint, which is defined as follows:

(7) If [... *T*...*C*...] is a contrastive answer to a question, then there is no alternative *T′* of *T* such that the speaker is willing to assert [...*T′*... *C* ...].

This constraint explains why *too* is obligatory in contexts like (8).

(8) a. A: What did Peter and Pia eat?
b. B: *Péter ate pàsta, and Pía ate pàsta.
c. B′: Péter ate pàsta, and Pía ate pasta, tòo.

The reasoning goes as follows. The first member of the answer 'Péter ate pàsta' is a partial answer to the question and therefore bears a topic accent. This triggers the implicature, through application of the distinctiveness constraint, that there is no alternative *α* to *Peter* such that the speaker is willing to assert '*α* ate pasta'. So, by the epistemic step usual in such reasoning, it follows that no one else but Peter ate pasta. The speaker cannot then resume his discourse with 'Pia ate pasta' without plainly contradicting himself. Krifka's proposal is that the semantics of *too* is such that it allows the violation of distinctiveness by explicitly stating a discourse relation. According to Krifka, *too* is stressed in such contexts, because it brings a strong assertion. As a side remark, Krifka notes that another way to answer the question in (8) would be to use a conjunction as in (9). In such a case, there is no contrastive topic accent, and the speaker conforms to the maxim of manner, by preferring (9b) over (8b).

(9) a. A: What did Peter and Pia eat?
b. B: Peter and Pia ate pàsta.

We note that Krikfa's reasoning relies crucially on the presence of a topic accent in the first part of the answer, and on the idea that this very accent triggers a distinctiveness implicature, which then has to be canceled via the use of the additive particle. The presuppositional nature of *too* (and of other additive particles) doesn't play any role.

2.1.3 *Sæbϕ*

The recent paper (Sæbϕ 2004) is directly concerned with the obligatoriness of *too*, and it brings several objections to Krifka's proposal. First, it is noted that there are contexts in which *too* is obligatory, even though there is no contrastive topic. It is the case, in particular, in narrative discourses like in (10).

(10) When the gods arrive at Jotunheim, the giants prepare the wedding feast. But during the feast, the bride —Thor, that is— devours an entire ox and eight salmon. He also drinks three barrels of beer. This astonishes Thrym. But Loki averts the danger by explaining that Freyja has been looking forward to coming to Jotunheim so much that she has not eaten for a week. When Thrym lifts the bridal veil to kiss the bride, he is startled to find himself looking into Thor's burning eyes. This time, (# Ø/too), Loki saves the situation, explaining that the bride has not slept for a week for longing for Jotunheim.

Sæbϕ also shows that even if one wants to take advantage of the presence of contrastive topics and the idea that they trigger a distinctiveness implicature, the computation should not be done from the first sentence, but from the second one. Thus, for instance, Krifka's reasoning does not explain why *too* is compulsory in an example like (11).

(11) Swift Deer could see pine-clad mountains on the other side of the Rain Valley. Far away to the east and west the dry prairies stretched out as far as the eye could see. (i) To the north lay the yellow-brown desert, a low belt of green cactus-covered ridges and distant blue mountain ranges with sharp peaks. (ii) To the south (# Ø/too) he could see mountains.

In this example, the speaker establishes a contrast between north and south. Let us assume that Krifka's analysis applies, and that there is a contrastive accent on *To the north*. Then it could be inferred, by application of the distinctiveness constraint on the sentence (i), that there is no alternative α such that the speaker would be willing to say that *to α lay the yellow-brown desert, a low belt of green cactus-covered ridges and distant blue mountain ranges with sharp peaks*. So, the constraint says that the speaker is not willing to assert, in particular, that *to the south lay the yellow-brown desert, a low belt of green cactus-covered ridges and distant blue mountain ranges with sharp peaks*. But this is not incompatible with what is said in the following sentence (ii). So there is no reason why *too* would be necessary, since there is no violation of the constraint.

So, Sæbϕ claims that to account for the obligatoriness of *too*, it is not necessary to bring in to play the presence of a contrast in the context, nor to appeal to a distinctiveness implicature. Rather, it is sufficient to analyze the proper meaning of *too*: adding the particle would introduce information meant to cancel an implicature that would otherwise be triggered by the sentence without *too*, and which

is in contradiction with the context. We won't present here the details of Sæbϕs analysis, but we retain two elements from it:

- first, the idea that the presuppositional character of *too* is more important than the fact that it associates with a contrastive topic;
- second, the idea that the reasoning takes as a starting point the implicatures and presuppositions triggered by the second sentence rather than by the first one.

Our proposal, relying on an implicature triggered, not by the presence of an accent in the first sentence, but by the possibility of use of *too* in the second one, is more in line with Sæbϕ's work than with Krifka's, and accounts for examples like (10) and (11).

2.1.4 *Intermediate conclusion*

What we take from the accounts very briefly summarized here is firstly that a proper account of the phenomenon has to take into account the fact that it is not limited to the well-known case of *too* ; a much larger class of particles (or presupposition triggers) exhibits the same behavior. Secondly, we also consider, after Sæbϕ, that even though contrast seems indeed to play a role, the presuppositional aspect (which has to do with discourse linking) should be investigated further.

This is why we try in the following sections to characterize as precisely as possible the class of items proving obligatory in the kind of contexts we have seen. But before doing so, we want to show why we consider that presupposition plays a bigger role than usually acknowledged.

2.2 The role of presupposition

2.2.1 *Discourse particles*

Zeevat's work (Zeevat 2002; Zeevat 2003) is also concerned with the class of obligatory items, proposing (among many other things) that the obligatoriness of *too* (and various other phenomena) be accounted for by considering a larger class of *discourse particles*, presupposition no longer playing a crucial role in the explanation. The argument relies in part on the observation that there is a set of particles that have in common (1) that they are obligatory (or, rather, not optional), (2) that they have a "minimal meaning", and (3) that they give rise to an accessibility anomaly. This class would contain *too*, but also particles like *indeed*.

It turns out that the class of triggers we want to consider have indeed the first two properties (see Section 2.3). As for accessibility, a few more words are necessary.

2.2.2 *Accessibility differences*

Zeevat proposes a list of obligatory triggers (Zeevat 2002, p. 85) which might serve as a starting point, but his list does not contain several presupposition triggers we want to consider, and contains particles which are not presuppositional and that we do not want to consider here. Let us start with these.

A good example is *indeed* (Zeevat 2003). We do not consider such a particle as triggering a presupposition, and we claim that there are two reasons why it should be kept separated from the triggers we consider.

Firstly, what Zeevat calls accessibility constraints do not apply identically for *indeed* and for *too*:

(12) a. *Mary dreamt that night that she would fail the exam and John will fail too.
b. Mary dreamt that night that she would fail the exam and indeed she did.

Here it is expected that *too* is not licensed because the antecedent is not accessible. *Indeed*, on the contrary, as a discourse particle, seems able to access the very same "antecedent".

Secondly, on the contrary, *too*, as a presupposition trigger, is not sensitive to embeddings that are presupposition holes, as in (13a–b).[2] As a consequence, *too* is obligatory even inside an embedding, as soon as its presupposition is satisfied. Compare with (13c).

(13) a. Jean est malade. Paul croit que Marie est malade (# Ø / aussi).
John is sick. Paul believes that Marie is sick (Ø / too).
b. Jean est malade. Est-ce que Marie est malade (# Ø / aussi) ?
John is sick. Is Marie sick (Ø / too)?
c. ?John is probably sick and Mary believes that he is indeed.

These examples seem to us harder to account for if it is considered that *too* (for example) is a discourse connective, since we expect discourse connectives to work differently when they are embedded (roughly, discourse connectives are sensitive to embedding, whereas presupposition triggers are not—leaving aside the well-known projection problem cases).

2.3 Generalization

Even though it may be the case that the phenomenon we are dealing with here is not limited to presupposition triggers, we still think that it is worth trying to define

2. It turns out that the embedding under negation of many of the triggers involved here cannot be done easily because they are polarity items (negative for *plus,* positive for *aussi, encore).*

precisely the sub-class of presupposition triggers that are obligatory, and that is what we try to do in this section.

2.3.1 *Inventory*

We have already seen many examples involving additive particles, and as is made clear by (Zeevat 2002), they all clearly prove obligatory:

(14) a. Jean est malade, Marie est malade (# Ø / aussi).
John is sick, Mary is sick (Ø / too).

b. Il était là hier, il est (# Ø / encore) là.
He was there yesterday, he is (Ø / still) there.

c. Paul est parti en Turquie l'an dernier, il ira (# Ø / de nouveau) cette année.
Paul went to Turkey last year, he will go (Ø / again) this year.

d. Jean est allé il y a deux ans au Canada. Il n'ira (# pas / plus) là-bas. =(la–b)
John went to Canada two years ago. He won't go there (Ø / anymore).

e. Léa a fait une bêtise. Elle ne la (# Ø /re-) fera pas. = (lc–d)
Lea did a silly thing. She won't (Ø / re-) do it.

The presuppositional complementizer of the (factive) verb *to know* exhibits the same behavior (even though it is harder to provide an appropriate context) (15). This trigger is usually not considered as additive. It should be noted that in French the class of (factive) verbs capable of introducing either a clause (with the complementizer *que*) or a question (with the complementizer *si*) seems to be very small, comprising in addition to *savoir* (*to know*) *ignorer* (*not to know*), *vérifier* (*check*), *comprendre* (*understand*) but not *découvrir* (*discover*), *réaliser* (*realize*)... In English the class is different (comprising *realize that vs. realize whether*, *be aware of/that vs. be aware whether*, at least for some speakers). We conjecture that the same obligatoriness can be shown for those verbs.

(15) a. [Léa est partie en Afrique.] Jean ne le dit à personne, bien qu'il sache (# si/que) elle est partie là-bas.
[*Lea's gone to Africa.*] *John tells no one, even though he knows. (whether/that) she's gone there.*

b. Jean est revenu de vacances. Mais comme il n'a téléphoné à personne, au bureau, tout le monde ignore (? si/que) il est chez lui.
John has come back from vacation. But since he called no one, at his office everybody 'ignores' (whether/that) he is at home.

c. Il y a eu une fuite d'eau, mais quelqu'un l'a réparée. Jean a appelé le plombier pour qu'il vérifie (? si/que) la fuite est réparée.
There was a leakage, but somebody fixed it. Jean called the plumber so that he checks (whether/that) the leak is fixed.

As for cleft constructions, data is more intricate, because in most situations where the presupposition associated with clefts is satisfied (e.g., in (16a)), it is very natural, at least in English, to use stress instead of a cleft construction. So, for instance, (16b) is quite acceptable and probably even more frequent than (16a).

(16) a. Someone fixed the dinner. It is John who did it.
b. Someone fixed the dinner. JOHN did it.

So, in a way, cleft constructions are not obligatory. But here what permits us to dispense of the cleft construction is the presence of another trigger: intonation, in such contexts, is usually considered to trigger a presupposition (in English at least) (Beaver 2001, p. 11, e.g.). Besides, the presupposition triggered by 'JOHN VP-ed' and by 'It is John who VP-ed' is the same. We conclude that what is compulsory is the use of one of the available triggers, which is confirmed by (17) where in the absence of any presupposition trigger, the discourse becomes deviant.

(17) #Someone fixed the dinner. John did it.

In French, it is not so clear that intonation behaves as a presupposition trigger, because of general properties of the French intonation system. For instance, in (18), there does not seem to be a way of stressing *Jean* that could render the example acceptable.

(18) a. Quelqu'un a préparé le dîner. Ce n est pas Jean qui l'a fait/# Jean ne l'a pas fait.
Someone fixed the dinner. It is not Jean who did it/Jean did not do it

In other cases, the behavior of French is closer to that of English:

(19) a. Quelqu'un a préparé le dîner. (C'est Jean qui/ JEAN / # Jean) l'a fait.
Someone fixed the dinner. (It is Jean who / JEAN / *Jean) did it*.
b. Paul n'a pas préparé le dîner. (C'est Jean qui / JEAN / # Jean) l'a fait.
Paul hasn't fixed the dinner. (It is Jean who / JEAN / *Jean) did it.*

So, we have to add to our inventory both cleft constructions and presuppositional intonation, which share the same presuppositional content, and which are such that when their presupposition is satisfied, it is obligatory to use one of them (so, in a way, obligatoriness is not attached to a specific lexical item or contruction, but rather to the set of available means to express one given presupposition).

Despite the large number of different triggers involved, it is clear however that all are not obligatory.

Consider for instance the trigger *regret*. For the sake of the argument, we can assume that (20a) presupposes (20b) and asserts (20c). If this trigger behaved similarly to the ones we have considered so far, then (20d) would be out, the only option being (20e). But both options are available.

(20) a. Bob regrets that it is raining.
b. It is raining.
c. Bob doesn't like it when it rains.
d. It is raining. Bob doesn't like it when it rains.
e. It is raining. Bob regrets that it's raining.

Similarly, the restriction trigger *only* can be analyzed as presupposing its prejacent (21b) and asserting the exclusion (21c). Then again, it is possible to form a discourse with both pieces of information without being forced to use the trigger (21d), while the version with the trigger is also possible (21e).

(21) a. Only Max owns a red car.
b. Max owns a red car.
c. No one else (than Max) owns a red car.
d. Max owns a red car, and no one else does.
e. Max owns a red car, and only Max does.

So we can't find situations where a trigger like *regret*, or *only*, is obligatory, and this is not really a surprise, because such triggers can't be added or removed without altering the asserted content of the host sentence.

2.3.2 *Definition of the class*

Let us try now to find a characteristic property of the triggers that give rise to this obligatoriness phenomenon. We take additivity, as Krifka defines it, as a starting point.

(22) **Additivity**

$[\text{ADD}[...F...]]: \underbrace{[...F...]}_{\text{asserted}} \underbrace{(\exists F' \neq F\ [...F'...])}_{\text{presupposed}}$ (Krifka 1999, p. 1)

An additive particle (ADD) is a particle such that when added to a proposition in which a constituent F is focused, it yields an interpretation that can be divided into two parts, an asserted content which is exactly the initial proposition (without ADD), and a presupposed content stating that there is an alternative F' such that replacing F with F' in the initial proposition gives a true proposition.[3]

It is quite easy to check that *aussi*, *non plus* (negative polarity version of *aussi*) fit with this definition (F' can be an individual, or a property, depending on what is focused).

3. In the case of *too,* which cannot freely accommodate, there might be an additional constraint: F' has to be given.

When it comes to other triggers usually considered as additive, like *encore*, *de nouveau*, *toujours*, the above definition has to be slightly amended. For instance, with *encore* (still), assuming an underlying event (*à la* (Davidson 1967)), or a state, or a time interval, the presupposed part would rather be something like $\exists F' < F$ [...F' ...], saying that there is another eventuality not only different from F but also temporally located before F in the past.

The previous definition assumes that we can easily separate the additive particle and the rest of the sentence (noted [... F ...]). When it comes to *ne... plus* in French, things become slightly more complicated, since simply removing *plus* would lead to an ungrammatical sentence (at least in modern French), and we have to admit that (23a) can be analyzed as being composed of a sentential negation (historically brought by *ne)* and an additive adverb (*plus (more)*). This seems reasonable, when considering other languages where the equivalent of *ne... plus* is a compound with a negation and an additive particle (*no more, nicht mehr, non piu...*).

(23) a. Il n'ira plus là-bas.
He won't go there anymore

Cleft constructions are much harder to analyze as additive in Krifka's sense: the problem comes from the presupposition, where there is no alternative (F') involved: if (24) was additive in the previous sense, the presupposed part would state that somebody other than Jean came, but it is not what a cleft sentence like this presupposes. On the contrary, it only presupposes that somebody came, and there is a strong tendency to pragmatically reinforce such a sentence to the reading that Jean is the only one who came. So cleft constructions do not fit the above definition of additive particles, unless one removes the difference condition (24b).

(24) a. C'est Jean qui est venu.
It is Jean who came.

b. [CLEFT [Jean$_F$ came]]: $\underbrace{[\text{Jean}(=F)\ \text{came}]}_{\text{asserted}}\ \underbrace{(\exists F' \neq F\,[F'\ \text{came}])}_{\text{presupposed}}$

For *to know whether/if*, it is intuitively easy to separate the asserted part from the presupposed part, but again, it is not possible to have it fit with Krifka's definition. To have it fit, we have to stipulate that *to know that* is a compound form : *to know whether + factivity*.

(25) a. Jean knows that it is raining

b. [FACT [Jean knows whether P]]: $\underbrace{[\text{Jean knows whether } P]}_{\text{asserted}}\ \underbrace{([P])}_{\text{presupposed}}$

So it is not possible to gather all the triggers here in the class of additive triggers, but they still have a common property, which can be stated as follows: their asserted content is reduced to their *clausal scope* (Krifka 1999), *i.e.* the clause to which they are added. As for their presupposed content, it is much more complex than in the definition (22) above.

(26) **Triggers with no asserted content**

$$\mathrm{TR}[\ldots F\ldots]]:\underbrace{[\ldots F\ldots]}_{\text{asserted}}\underbrace{[\ \ \Phi\ \ ]}_{\text{presupposed}}$$

The class of triggers with no asserted content comprises:

Pure additive items *aussi, non plus*. Presupposition: there is an alternative different from *F*.

Aspectual items *re-, encore, de nouveau, ne... plus*. Presupposition: there is an alternative eventuality located in the past.

Cleft and intonation Presupposition: the proposition is true of one member of an alternative set (not necessarily different from *F*).

Some factive verbs *savoir + que, ignorer + que, vérifier +que, realize, be aware*. Presupposition: the embedded proposition is true (factivity).

To sum up, this property defines a sub-class of presupposition triggers (not restricted to particles, neither to additive triggers) which are obligatory as soon as their conditions of use (presupposition) are (linguistically) satisfied.

3. Pragmatic explanation

We are looking for a general explanation, not relying on the semantics of a particular presupposition trigger, since the phenomenon involves an apparently heterogeneous class of triggers.

One common point in all our examples is that they somehow involve redundancy. Presupposition is a well-known tool for expressing redundancy (Roberts 1998, e.g.); we can thus characterize our examples as cases where redundancy is obligatory.

This is reminiscent of a classical test for presupposition, namely the fact that a proposition cannot be asserted after it has been presupposed, whereas the contrary (first an assertion, then a presupposition) is possible: see (27) (van der Sandt 1988, p. 161).

(27) a. Mary used to beat her husband. She has now stopped doing so.
b. #Mary has now stopped beating her husband. She used to beat him.

The usual explanation for this contrast has a pragmatic flavor: roughly, the speech act associated with assertion (bringing new information) cannot be felicitously

performed in a context where this information is already in the common ground. In contrast, the speech act associated with a presupposition is compatible (by definition) with a context where the presupposition is in the common ground.

This pragmatic flavor motivates the explanation we are elaborating in this section ; but it is worth noting that the phenomenon also interacts with discourse structure, which is dealt with in Section 4.

To explain the data, we can compare it with those studied in (Heim 1991) (we use Sauerland's (Sauerland 2003) presentation). Let us consider (28).

(28) a. #A wife of John's is intelligent.
b. The wife of John's is intelligent.
c. #A father of the victim arrived at the scene.
d. The father of the victim arrived at the scene.

Heim's proposal is to get inspiration from Hawkins' proposal: just like the classical scalar alternative set in (29a), which gives rise to the famous gricean quantity based implicature, the pair ⟨a, the⟩ forms a scalar alternative pair, when taking presupposition into account.

(29) **"Scalar alternatives"**
a. ⟨some, all⟩ assertion
b. ⟨a, the⟩ presupposition (Hawkins 1978)

More precisely, *the* bears more presuppositions (uniqueness presupposition) than *a*. Then the use of *a* implicates that the presuppositions of the other term of the scale are not satisfied (namely, that it is not true that John has only one wife, for the example (28a)), which is incompatible with world knowledge. This behavior is supposed to derive from a general principle labeled "maximize presupposition" in (Sauerland 2003): *make your contribution presuppose as much as possible* (see also (Percus 2006; Schlenker 2008)).

We use a similar method to explain our data, with the difference that the phenomenon at hand occurs in discourse and not in isolation. Let us consider the example given in (14a) (repeated in (30a)), where *aussi* is obligatory.

We start with the familiar scalar implicature computation. First, there is an (asymmetric) entailment relation given in (30b).[4] For expository reasons, we write this down as in (30c), where *A* stands for assertion, and *P* for presupposition.

4. This is true if, in a Russelian treatment of presupposition, we conjoin the assertive part and the presupposed part, as noted in (30c). This notation may be controversial, but we use it here for simplicity reasons and we believe that nothing important hinges on it.

We can then consider the two propositions in (30b) as forming a scalar alternative.[5] Then by a classical computation we get that by uttering A the speaker implicates that $(A \wedge P)$ is not appropriate (30d). An additional step is required: uttering A and implicating $\neg(A \wedge P)$ leads to the conclusion that the presupposition does not hold (30e).

(30) a. John is sick, Mary is sick too
b. Mary is sick too $\rightarrow$ Mary is sick
c. $(A \wedge P) \rightarrow A$
d. $A \rightsquigarrow \neg(A \wedge P)$
e. $\neg P$ = No one else than Mary (in the appropriate context) is sick

Now this implicature is in turn incompatible with the first part of the discourse (30a), namely, John is sick. So, it appears that the contrasts above can be predicted if sentences with and without presupposition triggers are considered as scalar alternatives. More precisely, we would have the following scales, with cases where the alternative is between the form with or without the trigger, and other cases where the alternative is really between two interchangeable items.

(31) a. ⟨pas (*not*), plus (*no more*)⟩
b. ⟨∅, aussi (*too*)⟩
c. ⟨si (*whether*), que (*that*)⟩

Let us come back to the comparison between our line of explanation and the one proposed in (Krifka 1999). Let us start with Krifka's computation. The sentence *John is sick* (let us assume we have the appropriate context in terms of topic/focus) will give rise to the distinctiveness implicature : 'no one else is sick'. Then uttering *Mary is sick* would result in a plain contradiction (leaving aside the fact that this second sentence would normally in turn give rise to a distinctiveness implicature). Then *too* has to be added, as a sort of reparation resort: it indicates that for some reason, the speaker doesn't want the distinctiveness implicature to go through.

Our proposal makes different assumptions. After the first sentence is uttered, the question of whether or not to add *too* arises (since it might form an alternative with not adding anything), but its presupposition is not satisfied, and so there is no real alternative, *too* is not added. Then the second sentence is uttered, and the alternative is once more considered. In this case, the presupposition of *too* is

5. Leaving aside the traditional problem of deciding why the two propositions are "natural" competitors.

satisfied in the context, so that the speaker really has a choice between adding the particle or not.

If the speaker chooses not to add *too*, then by the scalar reasoning given above, it would lead the hearer to infer that the speaker is reluctant to add too, thus that the presupposition is not satisfied, that is nobody else is sick, and this is contradictory with the context. Therefore, to prevent the hearer from making a contradictory implicature, the speaker is obliged to use *too*.

Our proposal accounts for several cases presented above, which are problematic for Krifka's account. Namely, the example in (3), where there is no identity between the first element (Bart is 17) and the second one (Wendy is old enough to have a driver's licence). Sæbϕ's examples are also accounted for. We can add the following very nice example, where *too* is not obligatory, but depending on its presence, two different readings are possible. From (32a) it is inferred that G. Romme is not Dutch, contrarily to what is inferred from (32b).

(32) a. The 5000 m race was won by Gianni Romme. The 1500 m race was won by a Dutch skater.
b. The 5000 m race was won by Gianni Romme. The 1500 m race was won by a Dutch skater too. (Sæbϕ 2004)

There remains several questions to be answered. First, one may ask where the "maximize presupposition" principle come from. It seems however reasonable to assume that it comes from Grice's (Grice 1975) maxim of quantity. The second question is harder to answer: how can we predict that the couple (s, s + *too*) forms a (scalar) alternative? In other words, why is a natural competitor of? We do not have a final answer to this question, but we think a lead worth considering is precisely the fact that all the triggers involved here have no asserted contribution. See (Zeevat 2002).

Finally, a question that deserves more development is the connection with accommodation of the presupposition. It is well known that, to put it in the words of van der Sandt and Geurts's (van der Sandt & Geurts 2001), "the presupposition of *too* is reluctant to accommodate" (see also (Zeevat 2003)). This property is not shared by all the triggers in our class (for instance, *again* seems to be quite capable of giving rise to accommodation) ; but it might help distinguish among the triggers we have grouped together, to get a better understanding of what makes them fit together. It is quite interesting to remark that, roughly, *too can* only be used when the presupposition is there (Zeevat 2003, p. 169), and it *has to* be used when the presupposition is there.

We turn now to discourse considerations, and set out in the next section how our proposal can be implemented within a Discourse Coherence perspective.

4. Interaction with discourse structure

It turns out that the obligatoriness of presupposition can be removed in some contexts, and we first describe relevant contexts in § 4.1, before trying to implement an analysis in the framework of SDRT (§ 4.2).

4.1 Discourse sensitivity

There are a number of apparent counter-examples that we have to deal with:

(33) a. Jean est malade, Marie est malade, Paul est malade, tout le monde est malade alors!
John is sick, Marie is sick, Paul is sick, everybody is sick then!

b. Il était là hier, il est là aujourd'hui.
He was there yesterday, he is there today.

These examples are fine, and no trigger seems necessary. What these examples have in common is that a discourse relation is readily available, *enumeration* in (a), and some sort of *contrast/parallel* in (b). This seems to be responsible for the non obligatoriness of the trigger. It should be noted, however, that the triggers are not forbidden either:

(34) a. Jean est malade, Marie aussi, Paul aussi, tout le monde est malade alors!
John is sick, Marie too, Paul too, everybody is sick then!

One could draw from this data the conclusion that the availability of a discourse relation somehow blocks the requirement of the principle advocated for above. But this is not the case: in the following examples, a discourse relation is explicitly stated (by the connective *c'est pourquoi*), without preventing the presupposition trigger from being obligatory.

(35) a. Jean est allé il y a deux ans au Canada. C'est pourquoi il n'ira plus là-bas.
John went to Canada two years ago. That's why he won't go there anymore

b. #Jean est allé il y a deux ans au Canada. C'est pourquoi il n'ira pas là-bas.
John went to Canada two years ago. That's why he won't go there

4.2 Preliminary "implementation" in SDRT

We want to make here an additional observation: in the following example, there are two similar triggers (re-and), each presupposing the same thing, and they are neither obligatory (only one is), nor forbidden.

(36) Lea a fait une bêtise. Elle ne la refera plus.
Lea made a mistake. She won't re-do it again

This is predicted by our principle: applied recursively, the principle compares the presupposed contents to tell whether we have an alternative. Once a trigger is inserted, the other alternative forms have exactly the same presupposed content, and thus they are not required, but they are not banned either, for it is expected that a presupposition trigger with no asserted content is licit as soon as its presupposition is satisfied.

So, what we'll try to implement now is a general principle which can be stated as in (37).

(37) A trigger (with no asserted content) is compulsory only if it brings strictly more satisfied presuppositions than the sentence without the trigger

Consider the discourse relation enumeration as in (33a). Roughly, our hypothesis is that an enumeration contour on the first sentence forces the second sentence *Mary is sick* to be linked to the context, in a way similar (if not identical) to what *too* would do. So, the trigger *too* does not bring strictly more presuppositions, and is therefore not required any more.

(38) *John is sick* + contour $\rightsquigarrow$ "Enumeration" $\rightsquigarrow \exists x(x = j \wedge \text{sick}(x))$ "cataphoric presupposition"

It is then possible to sketch an update rule which takes into account all that we've said earlier.

(39)
- When trying to attach a DRS K_β to a context K_τ:
- Let s be the sentence corresponding to K_β; let $\{a_1, a_2, \ldots a_k\}$ be the set of *presupposition triggers without asserted content* that can be adjoined to s.
- For each pair $(\emptyset, a_i)$, compare the number of satisfied presuppositions of the two members: to this effect, try to link/accommodate $psp(a_i)$ against the context K_τ via the usual procedure (Asher & Lascarides 1998).
- If – $psp(a_i)$ is satisfied, and
- there is a difference in the pair $\langle \emptyset, a_i \rangle$ in the number of presuppositions
- Then the choice of $\emptyset$ gives rise to the implicature that the presupposition is false (antipresupposition *à la* (Percus 2006)).

The predictions that we get are the following : in most cases, the sentence s has fewer presuppositions than $s + a_i$, so the principle applies. But what this rule predicts is that when, for any reason, the sentence s in itself already triggers a presupposition, then the principle no longer applies. This explains why it is not obligatory to use several triggers when they are available. This also explains what happens in *enumeration* cases : we consider that enumeration forces this forward link (be it presuppositional or not), so that the second sentence in the enumeration has to be linked to the context. Then there is no difference in the form with or without *too*.

As for example (33b), we can consider it a special case of enumeration, but we can also see it as a contrast, and then we have to provide the semantic definition of contrast with this forward link.

5. Conclusion

Starting from the obligatoriness of some presupposition triggers in discourse (when their presuppositions are satisfied in the context), we have shown that this phenomenon is not limited to additive particles, as has been previously assumed. We claim that obligatoriness defines a sub-class of presupposition triggers, characterized by the fact that they have no asserted content. The general explanation we provide then relies on a general pragmatic principle, which could be summarized as "maximize redundancy via presupposition binding". Finally, we have also tried to provide a general explanation to account for the fact that the obligatoriness of presupposition triggers seems to be sensitive to discourse relations.

References

Asher, N. & A. Lascarides (1998). The semantics and pragmatics of presupposition. *Journal of Semantics* 15, 239–299.

Beaver, D. (2001). *Presupposition and Assertion in Dynamic Semantics*. Studies in Logic, Language and Information. Stanford, CA: CSLI Publications.

Büring, D. (1998). *The 59th street bridge accent*. London: Routledge.

Davidson, D. (1967). The logical form of action sentences. In N. Resher (Ed.), *The Logic of Decision and Action*, pp. 81–95. Pittsburgh University Press.

Green, G.M. (1968). On *too* and *either*, and not just *too* and *either*, either. In *CLS (Chicago Linguistics Society)*, Volume 4, pp. 22–39.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts*, pp. 41–58. New York: Academic Press. Reprinted in (Grice 1989, pp. 22–40).

Grice, H.P. (1989). *Studies in the Way of Words*. Cambridge and London: Harvard University Press.

Hawkins, J.A. (1978). *Definiteness and Indefiniteness*: *A Study in Reference and Grammaticality Production*. London: Croom Helm.

Heim, I. (1991). Artikel und Definitheit. In A. von Stechow & D. Wunderlich (Eds.), *Semantik: Ein internationales Handbuch des zeitgenössischen Forschung*, pp. 487– 535. Berlin: de Gruyter.

Jackendoff, R.S. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge (Mass.): MIT Press.

Kamp, H. (2001). Presupposition computation and presupposition justification: One aspect of the interpretation of multi-sentence discourse. In M. Bras & L. Vieu (Eds.), *Semantics and Pragmatics of Discourse and Dialogue*: *Experimenting with current theories*. Elsevier.

Kaplan, J. (1984). Obligatory *too* in english. *Language* 60(3), 510–518.

Krifka, M. (1999). Additive particles under stress. In *Proceedings of SALT 8*, Cornell, pp. 111–128. CLC Publications.

Percus, O. (2006). Antipresuppositions. In U. Ueyama (Ed.), *Theoretical and Empirical Studies of Reference and Anaphora: Toward the establishment of generative grammar as empirical science*, pp. 52–73. Japan Society for the Promotion of Science. Report of the Grant-in-Aid for Scientific Research. Also available at Semantic Archive.

Roberts, C. (1998). Information structure in discourse: Towards an integrated formal theory of pragmatics. ms. The Ohio State University.

Sauerland, U. (2003, jun). Implicated presuppositions. Hand-out for a talk given at the *Polarity, Scalar Phenomena, Implicatures* Workshop, University of Milan Bicocca, Milan, Italy.

Schlenker, P. (2008). 'Be articulate': A pragmatic theory of presupposition projection. *Theoretical Linguistics* 34(3), 157–212.

Sæbø, K.J. (2004). Conversational contrast and conventional parallel: Topic implica-tures and additive presuppositions. *Journal of Semantics* 21(2), 199–217.

van der Sandt, R.A. (1988). *Context and Presupposition*. London: Croom Helm.

van der Sandt, R.A. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics* 9(4), 333–378.

van der Sandt, R.A. & B. Geurts (2001). Too. In *Proceedings of the 13th Amsterdam Colloquium*.

Zeevat, H. (2002). Explaining presupposition triggers. In K. van Deemter & R. Kibble (Eds.), *Information Sharing*, pp. 61–87. CSLI Publications.

Zeevat, H. (2003). Particles: Presupposition triggers, context markers or speech act markers. In R. Blutner & H. Zeevat (Eds.), *Optimality Theory and Pragmatics*, pp. 91–111. London: Palgrave-McMillan.

Conventionalized speech act formulae

From corpus findings to formalization*

Ann Copestake[a] & Marina Terkourafi[b]
[a]Computer Laboratory, University of Cambridge/[b]University of Illinois

This paper concerns the representation of formulae which conventionally encode particular illocutionary forces. Our aim is to provide an account of illocutionary force which allows the conventionalized formulae to be regarded as interpretive shortcuts. We propose an HPSG account in which the conventional illocutionary force of utterances is represented separately from their compositional semantics. The conventional illocutionary force does not replace part of the compositional interpretation (as it might on an idiom theory of speech acts) but instead adds to it. In this way, compositional semantics and conventional illocutionary force both remain available to the interpretation, and can, for instance, license dual responses.

1. Introduction

This paper concerns the representation of formulae which conventionally encode particular illocutionary forces.[1] The idea of a speech act formula may be intuitively illustrated with reference to well-known examples in English. In context, both (1) and (2) can be interpreted as requests to close a window, but (2) is intuitively more conventional/formulaic than (1).

(1) It's cold in here.
(2) Could you close the window?

*This research was supported by the UK Arts and Humanities Research Council (AHRC). Corresponding author: ann.copestake@cl.cam.ac.uk

1. Speech act formulae differ from indirect speech acts in that the former are defined as such on quantitative grounds (frequency counts), while the latter on semantic grounds (nonliterality). Thus, the two terms are not co-extensive, although there is considerable overlap between them: a direct speech act (e.g., an imperative used to perform a request) may still constitute a speech act formula relative to a context if it is the most frequent expression realizing a particular illocutionary force therein; conversely, of course, not all indirect speech acts are speech act formulae, but only those that meet the criterion of frequency relative to a context. For more on how frequency was assessed for the purposes of this analysis, see Section 3.

Our aim is to provide an account of illocutionary force which allows the conventionalized formulae to be regarded as interpretive shortcuts (as first described by Bach 1975, Morgan 1978).[2] That is, we assume that the use of a formula in a given context guides the hearer to a particular interpretation, but that the same interpretation could potentially have been reached by full inference about the speaker's desires and so on. The use of a conventional formula by a speaker can be assumed to make an intended illocutionary force clearer to the hearer, perhaps disambiguating intentions. Under this assumption, the conventional illocutionary force of utterances is represented separately from their compositional semantics. The conventional illocutionary force does not replace part of the compositional interpretation (as it might on an idiom theory of speech acts) but instead adds to it. In this way, compositional semantics and conventional illocutionary force both remain available to the interpretation, and can, for instance, license dual responses (cf. Clark 1979; Clark & Schunk 1980).

The remainder of this paper is organized in six sections. In Section 2, we present the empirical motivation for our approach, which builds on a conversational corpus of Cypriot Greek offers and requests recorded in a variety of settings. Analysis of these data revealed that realizations of offers and requests clustered around different verbal expressions, depending on the situational context, prompting definition of these expressions as speech act formulae. Five such formulae are presented in Section 3. Section 4 outlines our proposal for representing these formulae in Head-driven Phrase Structure Grammar (HPSG), while Section 5 discusses some advantages of this proposal over alternative ones. Finally, Section 6 summarizes the argument and suggests some directions for future research.

2. Empirical motivation

Our work is based on a corpus of 2,189 spontaneous exchanges in Cypriot Greek (Terkourafi 2001). Conversations between native speakers were tape-recorded in various settings (at home, at work, on radio/TV) and later transcribed. In this way, several physical or otherwise extra-linguistic features of the situation, such as the gender, age, social class of interlocutors, the relationship between them, and the setting of the exchange were available and independently noted. Later,

2. Our use of the term 'conventionalized' springs exactly from the probabilistic character of Morgan's 'conventions of usage', which remain defeasible, contrary to his 'conventions of the language' that are no longer so. Of course, these two types of conventions are not unrelated, the former feeding regularly into the latter. Although in what follows we use the terms 'conventionalized' and 'conventional' interchangeably, we always have in mind conventions of usage, that are pragmatic (defeasible) and not semantic in nature.

these features served to reduce the fully actualized, nonce contexts of occurrence to schematic or otherwise 'minimal' contexts consisting of the values of a limited set of contextual parameters (e.g., men, aged 31–50, of middle class, addressing women, aged 31–50, of middle class, for the first time, in a relationship of new customer to salesperson in a shop) but with all other specificities removed. Analysis of these data aimed to establish the linguistic means by which offers and requests are realized in each minimal context, and whether these linguistic means can be better predicted by a general rationality principle,[3] or with reference to particular minimal contexts with which they are directly related by culture-specific convention.

Definitions of offers and requests commonly appeal to the dimensions of speaker vs. hearer agency (of whom the act is predicated), or hearer vs. speaker benefit (who benefits from the predicated act) respectively. However, neither of these is unproblematic. In many contexts, an activity involves several agents cooperating to achieve a mutually beneficial outcome. For instance, a transaction in a shop can be viewed as a buying event, in which case the customer is the agent, or as a selling event, in which case the agent is the shopkeeper. To overcome these difficulties, during the analysis, utterances realizing offers or requests were selected and classified as such depending on the addressee's uptake (Austin 1962). In other words, our working definition of a request for the purposes of data analysis was very much a pragmatic one ('what is responded to as a request, counts as a request'), and *mutatis mutandis* for offers. This was possible due to the methodology adopted during collection of the data, and safeguarded against the circularity that might result from analyst bias in associating particular constructions with particular speech acts. When uptake was unavailable or otherwise insufficient, desirability to speaker/hearer assessed based on the propositional content of the utterance was used as a supplementary criterion (Terkourafi 2001: 33–36, 39–44). On this view, (3) is an offer (to fetch a blanket), despite the fact that the agent of the action explicitly mentioned is the hearer. We return to such examples in 4.1 below.

(3) [At home; Speaker: female, aged over 51, working class; Addressee: female, aged over 51, middle class; Relationship: friends]

thelis na skepastis?
want-2SG SUBJ cover-PASS-2SG?[4]
'Do you want to cover up?'

3. Brown and Levinson's (1987: 76) formula $W_{FTAx} = D_{istance}\,(S,H) + P_{ower}(H,S) + R_{anking\,FTAx}$) was used for this purpose.

4. Transcription conventions: SUBJ=subjunctive particle; FUT= future particle; NEG=negative particle; PASS=passive; ?=rising intonation; (.)=brief pause; .hh=audible inhaling; (( ))=transcriber comment.

Quantitative and qualitative analysis of the linguistic realizations of offers and requests in different minimal contexts revealed the existence of formulae used to perform these acts in particular minimal contexts. We are dealing with a speech act formula when, given a minimal context and a set of semantically equivalent expressions conveying the same illocutionary force, one expression is preferred above others to realize offers or requests in this context. This is not the same as a particular pragmatic strategy being preferred, since each pragmatic strategy can be realized by several linguistic expressions (Brown & Levinson 1978/1987). That is, the association between speech act formula and illocutionary force in a minimal context turned out to be more specific than the one between pragmatic strategy and context, and also conventional, such that it could not be predicted by a general rationality principle.[5] Moreover, the distribution of formulae was affected by extra-linguistic features: speaker's and hearer's gender, age and social class, the relationship between them, the setting of the exchange, and the sequential placement of the utterance in the discourse.

This survey is probably the most extensive one to date for any language where the full context of natural interactions was directly observed (as opposed to studies on previously collected corpus data or purpose-designed experiments involving artificial tasks and/or role-play). We believe these corpus findings provide grounds for an alternative analysis of illocutionary force as a phenomenon that is not exclusively a matter of full-fledged inference about individual speakers' intentions but is partly predictable on the basis of contextual features such as those that co-constitute a minimal context.

The theoretical background for our proposal is provided by Morgan's (1978) notion of 'conventions of usage.' The problem is, of course, how to define a convention of usage empirically, and how to go about describing it in a way that generalizes across conventions. Previous efforts to deal with this problem include plan-based approaches (e.g., Cohen & Perrault 1979; Allen & Perrault 1980; Perrault & Allen 1980). Our attempt to generalize over speakers' intentions builds on these approaches.[6] At the same time, the Cypriot Greek corpus findings allow us to go beyond them, in empirically grounding the notion of convention of usage to features of context whose values are immediately observable; or, at least, treated as such, in that they are often settled by presumption (Terkourafi in press).

---

5. See Terkourafi (2002: 184–192) for a proposal of how to account for this relationship from a diachronic, macro-social perspective.

6. Terkourafi (in press) draws an explicit connection between minimal contexts and plans. Specifically, she suggests that the contextual features that jointly constitute a minimal context can serve as short-cuts to the intended illocutionary force because they are actually 'plan-enabling' features for plans involving particular illocutionary forces.

3. Cypriot Greek speech act formulae

Most of the formulae emerging from the analysis of the Cypriot Greek data outlined in the previous section are grounded in particular lexemes, especially inflected verbs rendered with a particular accent and intonation.[7] However, grammatical constructions such as imperative or 1st person singular subjunctive verb-forms may also receive a formulaic interpretation. The main criterion for identifying a formula is frequency in a 'minimal' context as outlined above (i.e., a set of co-occurring extra-linguistic features) but evidence of lexicalization, including a fixed word order, phonological reduction and a characteristic intonation contour, is also taken into consideration.

We illustrate the notion of a speech act formula with reference to five such formulae that emerged from the analysis of the Cypriot Greek corpus data: *echete NP?* (have-2pl NP?), *eʃi NP?* (have-3sg NP?), *tha ithela VP* (FUT want-PAST-1sg VP), *thelo VP* (want-1sg VP), and *thelis NP/VP?* (want-2sg NP/VP?). The first four almost always realize requests, while the last one almost always realizes offers. Use of the first formula is exemplified in (4).

(4) [In a pub; Speaker: male, aged 31–50, middle-class; Addressee: female, aged 18–30, working-class; Relationship: new customer to salesperson]
echete: pʰinats?
have-2PL peanuts?
'Do you have some peanuts?'

When uttered as an opening request by middle-class customers who walk into a shop for the first time addressing a service-provider, *echete NP?* typically realizes a request. This is its most frequent interpretation in the data, accounting for six out of a total of nine occurrences of *echete NP?* in the corpus. Moreover, compared with other verb-forms, *echete NP?* is the most frequent one used in middle-class first-time customers' opening requests addressed to salespersons: it is used 15% of the time in this context, with other verb-forms following at 12% of the time or less.

The second formula, *eʃi NP?*, is similarly preferentially associated with opening requests by first-time customers addressing salespersons, but this time the customers are working-class, as in (5).

7. Given the non-standard character of Cypriot Greek compared with Standard Modern Greek, it is necessary to differentiate broadly (regional) 'accent' (concerning mainly phonetic realization) from 'intonation' (concerning mainly prosodic features of the utterance).

(5) [In the open-air market; Speaker: female, aged 31–50, working-class; Addressee: male, aged 31–50, working-class; Relationship: new customer to salesperson]
eʃi mikres pu na min echun scheδia pano? aspro
have-3SG small that SUBJ NEG have patterns on? white
'Are there any small plain ones? In white.'

Again, this is the most frequent interpretation of *eʃi NP*? in the data, accounting for twenty out of its twenty-three total occurrences in the corpus. *eʃi NP*? is also the most frequent verb-form in opening requests from working-class first-time customers to salespersons, being used 15% of the time.

Contrary to the 'commercial transaction' frame in which the previous two formulae were used, the following two formulae prevailed in formal discussions on the radio and television, as in (6) and (7).

(6) [On TV; Speaker: female, aged over 51, middle-class; Addressee: male, aged 31–50, middle-class; Relationship: interviewee to interviewer]
tha ithela na prostheso kati edho omos
FUT want-PAST-1SG SUBJ add-1SG something here though
'I would like to add something here though.'

(7) [On TV; Speaker: male, 31–50, middle-class; Addressee: male, aged 31–50, middle-class; Relationship: interviewer to interviewee]
thelo na mbume sto thema ton piravlon
want-1SG SUBJ go into-the subject of-the missiles
'I want us to come to the question of the missiles.'

In these formal settings, *tha iθela VP* is typically used to perform requests by interviewees addressing interviewers. This is its most frequent interpretation in the data, accounting for twenty-four out of its twenty-six total occurrences in the corpus. *tha ithela VP* is also the most frequent verb-form realizing requests by interviewees to interviewers being used 18% of the time, with other verb forms following at 13% of the time or less. *thelo VP*, on the other hand, is typical of requests addressed by interviewers to interviewees. This is its single most frequent interpretation in the data, accounting for sixteen out of a total thirty-seven occurrences, the remaining twenty-one occurrences being distributed among several other interpretations. *thelo VP* is also the most frequent verb-form used by interviewers in requests addressed to interviewees, being used 22% of the time, with other verb forms following at 18% of the time or less.

The final formula, *thelis NP/VP?*, is independently preferred across a wide range of informal contexts to realize offers. Moreover, its interpretation as a commissive is the most frequent one, accounting for 103 out of a total of 112 occurrences in the corpus as a whole. For this formula there is direct evidence

of lexicalization, for instance, the reduced phonology indicated as *'lis* in (8). In addition, *thelis NP/VP?* occurs utterance-initially in over 90% of instances in the corpus, the only items that can precede it being address terms, or the conjunction *lipon* ('so'):

(8) [In a shoe-shop; Speaker: female, aged 18–30, working-class; Addressee: female, aged 18–30, middle class; Relationship: acquaintances]

'lis kafe? (.) indalos in' o kafes su?
want-2SG coffee? How is the coffee your?
'Do you want coffee? How do you take your coffee?'

In addition to frequency relative to a context and evidence of lexicalization, a standard argument for treating some English request forms as conventional is the different distribution of IFIDs (Illocutionary Force Indicating Devices) such as *please* with requests such as (1) and (2) above (Green 1975). This argument extends to the corresponding Greek forms *parakalo* (formal 'please') and *ligho* (lit. 'a little', functioning as an informal variant of *parakalo*; cf. Sifianou 1992: 168). Whereas these are acceptable with conventional requestive formulae such as those in (4)–(7) above and in (9) below, they do not sound natural, and thus are not found, with non-conventional requests such as (10):

(9) [On the radio; Speaker: male, 31–50, middle-class; Addressee: male, aged over 50, middle-class; Relationship: interviewer to interviewee]

... tha ithela na mas pite tora na erthume ligho .hh eh sta tu iku mas
... FUT want-PAST-1SG SUBJ us tell-2PL now SUBJ come-1PL a little er to-the of-the house our
… 'I'd like you to tell us now, to come for a moment to our internal affairs.'

(10) [At a meeting of the philatelic society; Speaker A: female, aged over 51, middle class; Speaker B: male, aged over 51, middle class; Relationship: old colleagues]

A: *e niko? echo edho ta eksodha tis italias* (.)
B: *ne*

A: hey Nick? have-1SG here the expenses of-the Italy
B: yes

A: 'Hey Nick? I have the expenses from Italy with me.'
B: 'Yes.'

(where A's turn was interpreted by B as a request to reimburse the expenses)

4. The specification of formulae in HPSG

The verbal formulae considered above all correspond to inflected forms (with certain subcategorization) rather than lexemes. Thus what must be represented

in an HPSG account is a conventional association between an illocutionary force specification and a sign which has normal syntax and semantics, a particular intonation (important since interrogativity in Modern Greek is signaled prosodically), and possibly reduced phonology.[8] For instance, the sign involved in the *echete NP?* formula could be generated from the lexeme for *echo* ('to have') by applying the lexical rules for 2nd person plural and for rising intonation. The exact specification of the syntax of the signs is not important here, since there is nothing unusual about them. What we do have to consider is the specification of the context features and the status of the stipulation that relates this to the rest of the sign, which we outline briefly here.

The simplest option is to treat the formulae as analogous to lexical entries (or idioms) in that: (a) each formula is a conventional association between phonology, syntax, compositional semantics and context and, (b) all formulae are listed. Information specified about the formulae may further include their intonational contour, as this is a potential indicator of function (Rodriguez & Schlangen 2004) and situational context (Wichmann 2004). One possible hypothesis is that interrogatives conventionalized for a requestive function will be realized without the final rise of genuine questions. This hypothesis stems from previous research associating the conduciveness of a question (or, in different terms, the relative lack of ambiguity of an impositive interpretation for interrogative directives) with a fall-to-low terminal tone (Brown et al. 1980: 186–7; Bartels 1999: 267–273). Though this hypothesis remains to be fully investigated, preliminary experimental results suggest its plausibility (Nickerson & Chu-Carroll 1999). In the remainder of this section we will consider the interaction between conventional illocutionary force and compositional semantics.

We assume that a C-ILLOC feature in CONTEXT is used to represent conventional illocutionary force. This will only be specified in utterances where a formula is used. We remain neutral about whether there is any representation of inferred speaker intentions in the sign.[9] We can thus distinguish three possibilities for utterances:

a. No conventional illocutionary force. Speaker intentions must be derived by the hearer (via inference) from compositional semantics.

---

8. We do not consider examples of formulae with non-standard syntax and semantics in this paper, but see Terkourafi (2009) for a preliminary analysis.

9. A feature structure could be used to represent such intentions, even if the formalism is not well-suited to computing them, so it would be possible to assume some external component, analogous to the way morphophonology is indicated in Pollard and Sag (1994).

b. Conventionalized illocutionary force, C-ILLOC, is instantiated along with compositional semantics. The hearer assumes that the speaker intentions are given by C-ILLOC (defeasibly).
c. C-ILLOC is instantiated, but there is no (useful) compositional semantics.

We will now discuss each of these classes in turn, concentrating on the second, which is the focus of this paper.

The first class of utterance, where no illocutionary force is conventionally specified, is exemplified by (11)–(13). On our account, C-ILLOC is only instantiated if a formula is involved. It is not present in (11), for instance, even though the compositional semantics directly indicates that a request is being made, because this is not a use of a conventional formula. The inference from this to the speaker's intentions is presumably simpler than it is in (1) or (10), since the intention follows directly from the verb semantics, but the difference is not one that need be explicitly encoded.

(11) [Discussion in parliament; source: http://www.parliament.cy/parliamentgr/010/Documents/Vouli%20Geronton_praktiko-6Nov2004.doc; accessed 5/18/09]

Zito ke parakalo na ghini anavathmisi tu kendru
Ask-1SG and entreat-1SG SUBJ be-done upgrading of-the centre.
'I ask and petition that the centre be upgraded.'

C-ILLOC is similarly not instantiated for (12) and (13) but in these cases, rather than directly indicating illocutionary force, compositional semantics provides the input to a process of inference about speaker intentions, yielding their interpretations as a request and as an offer respectively.

(12) [At home; Speaker A: female, aged over 51, working class; Speaker B: female, aged 31–50, middle class; Relationship: mother to daughter]

A: *zina to moron pai potʃi*
B: *ne*

A: Zina the baby go-3SG from-there
B: yes

A: 'Zina, the baby is going to the other room.'
B: 'Yes.'

(13) [In an office; Speaker A: female, aged 31–50, middle class; Speaker B: male, aged 31–50, middle class; Relationship: employee to employer]

A: i*rthen efimeridha kirie* ((first name))
B: *'ndaksi mbravo egho mja mathkja* ((unintelligible))

A: come-PAST-3SG newspaper Mr ((first name))
B: OK bravo I one look ((unintelligible))

A: 'The newspaper has arrived Mr ((first name)).'
B: 'OK good. I ((just wanna take?)) a look.'

In the second class of utterance, illocutionary force is specified by a convention of usage and C-ILLOC is instantiated along with the compositional semantics. The values of C-ILLOC we consider are REQUEST and OFFER which can be formalized along the lines of Perrault and Allen's (1980) approach. There, for instance, REQUEST(S,H,ACT) (where S is Speaker, H is Hearer and ACT is some action) has the constraint that H is the agent of ACT, the precondition that WANT(S,ACT(H)) (the speaker wants the hearer to perform some act), the body BELIEVE(H,WANT(S,ACT(H))) (the hearer believes that the speaker wants them to perform the act) and the effect that WANT(H,ACT(H)) (the hearer wants to perform the act). The nature of the act will generally be specified by the compositional semantics.

As a concrete example, take the case of *thelis* ('do you want?') with a VP argument where the speaker is the agent of the VP, as in (14).

(14) [At a shoe-shop; Speaker: female, aged 18–30, working class; Addressee: female, 31–50, working class; Relationship: salesperson to new customer]
thelis na valumen kanena pataki mesa?
want-2SG SUBJ put-1PL any insole in?
'Do you want us to put an insole in?'

C-ILLOC is specified as an OFFER in the formula. The ACT which is being OFFERed is given by the second argument to the relation supplied by *thelis* in the compositional semantics (CONTENT). Schematically:

CONTENT: INT(want (H, 1 put-insole-in(S)))
C-ILLOC: OFFER(S,H, 1)

In (14), the action offered is directly specified from the compositional semantics. However, in a subset of the examples of conventional formulae this direct link is not possible. A relatively straightforward example is (8) where the argument to *thelis?* is an NP. Here, the interpretation involves conventionalized metonymy and the argument to OFFER could be fleshed out as OFFER(S,H,TRANSFER(S,H,NP)), where TRANSFER is some suitably generic predicate.

A more complex case was shown in (3). In context, (3) realizes as an offer, but our formal OFFER requires that the agent of the action be the speaker. Thus the C-ILLOC of (3) cannot be OFFER(S,H,cover-up(H)). To allow for examples such as (3), we have to distinguish between the cases where *thelis?* takes a VP with the speaker as the subject, which can be construed directly as conventional offers, and the cases where the subject of the complement VP is the hearer, which we treat as

conventional offers with an indirection between OFFER and the specified action. Thus the C-ILLOC of (3) corresponds to something like:

OFFER(S,H,ACT′(S)) ^ PRECONDITION(ACT′(S),cover-up(H))

That is, there is some implicit action ACT′ which is offered, which is a precondition to the explicitly mentioned action of the hearer.

The third class of utterances mentioned above covers the case where C-ILLOC is instantiated but there is no (useful) compositional semantics. Greetings such as *Hello!* are examples of this. However it is not clear whether requests or offers ever fall into this category.

We now turn to the issue of how C-ILLOC values are built up in the HPSG analysis of utterances. The value of C-ILLOC on a phrase is taken to be the unification of the C-ILLOC values of the daughters. Only a single conventional illocutionary force can be specified on each sign and thus if more than one daughter has a C-ILLOC value, the values would have to be mutually consistent. For instance, in (9) above, two C-ILLOC values are instantiated, one conventionally associated with the formula *tha ithela* and another due to the lexical item *ligho* (informal 'please'). Since both are specified as C-ILLOC REQUEST, unification succeeds.

In contrast, if *parakalo* (formal 'please') with a C-ILLOC value of REQUEST were to be added to (3) above, which conventionally bears the C-ILLOC value of OFFER due to use of the formula *thelis VP?*, the values would be incompatible and the utterance would not receive an analysis as an OFFER. However, a putative utterance such as (15) could receive an interpretation as a request given appropriate contextual support.

(15) *thelis na skepastis parakalo?*
want-2SG SUBJ cover-PASS-2SG please?
'Do you want to cover up please?'

We assume that this is because in (15) *thelis VP?* is not used as a formula, but has a purely compositional interpretation. In general, we assume that a non-formulaic interpretation is available when the formulaic interpretation is blocked by a conflicting C-ILLOC value, but leave a detailed account of this to future work.

5. Dual uptake

A potential advantage of representing C-ILLOC in CONTEXT and separately from compositional semantics is that the compositional interpretation is available to license dual responses (cf. Clark 1979; Clark & Schunk 1980). For instance, in response to A's turn in (16), B provides an answer to the question as well as an action complying with her request.

(16) [At a pharmacy; Speaker A: female, aged 18–30, working class; Speaker B: female, aged over 51, middle class; Relationship: new customer to salesperson]
A: *na mu kopsete apodhiksin?*
B: *ne* ((issues receipt))
A: SUBJ me cut-2PL receipt?
B: yes ((issues receipt))
A: 'Can you give me a receipt?'
B: 'Yes.' ((issues receipt))

The two parts of B's response in (16) are not equivalent. This can be seen in two ways. First, in the ordering of the two parts: responding to the interrogative semantics of A's turn takes place *before* complying with her request (or possibly overlaps with the beginning of the action). The reverse order of responding (first to the request, and then to the question) is intuitively ungrammatical, and indeed all 79 dual responses found in the corpus exhibit the order seen in (16), i.e., the answer always precedes compliance. The second way in which lack of equivalence between the two parts of B's response shows up is in the interactional consequences of providing only one of them. Not providing a verbal answer to the interrogative, while complying with the request, may at worse result in perceived impoliteness (cf. Clark & Schunk 1980).[10] However, merely providing an answer but not proceeding to comply with the request can be considered downright uncooperative, and the 'smart alecky' nature of such responses has been noted more than once (cf. Bach & Harnish 1979; Bertolet 1994; Terkourafi 2001).

Other authors consider dual uptake. Asher and Lascarides (2001: 193ff.) motivate the introduction of dot types for conventionalized indirect speech acts by dual uptake but they do not provide an account of the asymmetry noted above. Representing the compositional semantics of utterances separately from their illocutionary force in our approach may enable us to capture this basic asymmetry.

Our approach is in contrast to Terkourafi and Villavicencio (2003), who assume that a conventional formula gives a default illocutionary force in the feature structure which may be overridden by an inferred value. However, *please*-insertion etc, depend on the utterance being a conventional request, a distinction which they cannot capture. Furthermore, if the value is a default, there is little predictive power in the account, since in any constraint-based approach to defaults there can be no penalty for overriding a default. Finally, the overriding account implies a single computed illocutionary force and gives no insight into dual responses.

---

10. Clark (1980) reports on a series of experiments showing that the most important consequences of seriously attending to the literal meaning of indirect requests (by, e.g., acknowledging it in one's reply) lie with politeness.

6. Conclusion and future work

We have outlined an approach to encoding conventional speech acts as conventions of usage within HPSG that is grounded on extensive empirical data. Our account is intended to capture Morgan's (1978) insight that conventional speech acts are interpretive shortcuts. We believe that Morgan's argumentation for this type of approach is still essentially valid. Although modern versions of the performative hypothesis, such as Ginzburg and Sag (2001), Ginzburg et al. (2003), could perhaps be extended to the formulae we have considered, it is not at all clear that such an approach is sufficiently flexible to account for the observations we have discussed.

However, our account remains somewhat schematic. Although Green (2000) discusses a wide range of issues that affect the treatment of CONTEXT in HPSG, overall the literature is limited and this makes it difficult to refine our assumptions about C-ILLOC, for instance. One important aspect of Terkourafi's work is the finding that the use of formulae is heavily dependent on extra-linguistic features (as described in §2). Terkourafi and Villavicencio (2003) give a formalization of this in terms of a set of features in BACKGROUND. The intention there seems to be to hardwire these features as part of the formulae, but an alternative is to regard the distribution as more probabilistic in nature. This point is particularly relevant to the defeasibility of intentions inferred from the conventionalized C-ILLOC of speech act formulae: empirically, conventionalization shows up in relation to particular minimal contexts and not when considering the corpus as a whole, which means that there are other contexts in which a particular speech act 'formula' may not function as a formula (i.e., as an interpretive shortcut) at all but will always require full-blown reasoning instead, leading to an open-ended list of potential interpretations.[11] Future work will be aimed at fleshing out this account, by looking at data from English and Japanese in addition to the Cypriot Greek data and by considering discourse particles in more detail, as well as refining the encoding of the background features for extra-linguistic information.

References

Allen, James F. & Raymond C. Perrault. 1980. "Analyzing intention in dialogues." *Artificial Intelligence*, 15 (3): 143–178.

Asher, Nicholas & Alex Lascarides. 2001. "Indirect speech acts." *Synthese* 128: 183–228.

Austin, John. 1962. *How to Do Things with Words.* Oxford: Clarendon.

11. These two possibilities are captured with reference to the notions of generalized and particularized conversational implicatures in Terkourafi 2003.

Bach, Kent. 1975. "Performatives are statements too." *Philosophical Studies* 28 (4): 229–236.
Bach, Kent & Robert M. Harnish. 1979. *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press.
Bartels, Christine. 1999. *The Intonation of English Statements and Questions: A Compositional Interpretation*. New York & London: Garland.
Bertolet, Rod. 1994. "Are there indirect speech acts?" In *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*, Savas Tsohatzidis (ed.), 335–349. London: Routledge.
Brown, Gillian, Karen Currie, & Joanne Kenworthy. 1980. *Questions of Intonation*. Baltimore: University Park Press.
Brown, Penelope & Stephen C. Levinson. 1978/1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
Clark, Herbert H. 1979. "Responding to indirect speech acts." *Cognitive Psychology* 11: 430–477.
Clark, Herbert H. & Dale H. Schunk. 1980. "Polite responses to polite requests." *Cognition* 8: 111–143.
Cohen, Paul R. & Raymond C. Perrault. 1979. "Elements of a plan-based theory of speech acts." *Cognitive Science* 3: 177–212.
Ginzburg, Jonathan & Ivan Sag. 2001. *Interrogative Investigations: The Form, Meaning and Use of English Interrogatives*. Stanford: CSLI.
Ginzburg, Jonathan, Ivan Sag & Matthew Purver. 2003. "Integrating conversational move types in the grammar of conversation." In: *Perspectives on Dialogue in the New Millennium*, Peter Kühnlein, Hannes Rieser & Henk Zeevat (eds.), 25–42. Amsterdam: John Benjamins.
Green, Georgia M. 1975. "How to get people to do things with words." *Syntax and Semantics* 3, 107–141.
Green, Georgia M. 2000. "The nature of pragmatic information." In: *Grammatical Interfaces in HPSG*, Ronnie Cann, Claire Grover & Philip Miller (eds.), 113–138. Stanford: CSLI Publications.
Gunlogson, Christine. 2003. *True to Form: Rising and Falling Declaratives as Questions in English*. New York: Routledge.
Morgan, Jerry. 1978. "Two types of convention in indirect speech acts." In: *Syntax and semantics, vol. 9: Pragmatics*, Cole, Peter (ed.), 261–280. New York: Academic Press.
Nickerson, Jill. & Jennifer Chu-Carroll. 1999. "Acoustic-prosodic disambiguation of direct and indirect speech acts." *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, California.
Perrault, C. Raymond & James F. Allen. 1980. "A plan-based analysis of indirect speech acts." *American Journal of Computational Linguistics* 6: 167–182.
Pollard, Carl & Ivan Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
Rodriguez, Kepa J. & David Schlangen. 2004. "Form, intonation and function of clarification requests in German task-oriented spoken dialogues." *Proceedings of Catalog'04 (The 8th Workshop on the Semantics and Pragmatics of Dialogue, SemDial04),* Barcelona, Spain.
Terkourafi, Marina. 2001. *Politeness in Cypriot Greek: A frame based approach*. Ph.D. Thesis. University of Cambridge, Department of Linguistics.
Terkourafi, Marina. 2002. "Politeness and formulaicity: evidence from Cypriot Greek." *Journal of Greek Linguistics* 3: 179–201.

Terkourafi, Marina. 2003. "Generalised and particularised implicatures of linguistic politeness." *Perspectives on Dialogue in the New Millennium*, Kühnlein, Peter, Rieser Hannes & Henk Zeevat (eds.), 149–164. Amsterdam: John Benjamins.

Terkourafi, Marina. 2009. "oi na+V in Cypriot Greek: A speech-act construction at the interface of semantics, pragmatics and intonation." Presented at *Frames and Constructions: A conference in honor of Charles J. Fillmore, Aug 1, 2009, Berkeley, California*.

Terkourafi, Marina. in press. "On de-limiting context." In *Context in Construction Grammar*, Alexander Bergs & Gabriele Diewald (eds.), Amsterdam: John Benjamins.

Terkourafi, Marina & Aline Villavicencio. 2003. "Toward a formalization of speech act functions of questions in conversation." In *Questions and Answers: Theoretical and Applied Perspectives*¸ Raffaella Bernardi & Michael Moortgat (eds.), 108–119. Utrecht Institute of Linguistics OTS.

Wichmann, Anne. 2004. "The intonation of please-requests: A corpus-based study." *Journal of Pragmatics* 36: 1521–1549.

Constraints on metalinguistic anaphora*

Philippe De Brabanter
Institut Jean Nicod, Université Paris 4-Sorbonne

Just as it is possible to refer to any entity, concrete or abstract, in extralinguistic reality, it is also possible to refer to any linguistic entity, be that a phoneme, a word, a sentence, or any other linguistic object. Metalinguistic reference can be achieved in two ways, by means of 'autonymous' and 'heteronymous' mention (cf. Recanati 2000: 137). Autonymous mention is a matter of quotation: a token of a linguistic string can be produced in order to refer to another token of that string or to a type which it instantiates:

(1) "**Boston**" is disyllabic. (Quine 1940: 26)

(2) She said "**why don't you just drop dead?**"

Heteronymous mention concerns all the non-quotational means that can be resorted to in order to refer metalinguistically, namely descriptions and various types of pronouns.

(3) **The 4354th word of Chants Democratic** is disyllabic. (Quine 1940: 26)

The central difference between autonymous and heteronymous mention is that the former rests on an iconic relation between the mentioning expression and its linguistic referent, whereas the latter involves no such resemblance at all.

The focus of this paper is on a subset of heteronymous mention, namely those cases in which the mentioning expression is, roughly speaking, anaphorically linked to the string it mentions. I will distinguish two subclasses. In the first one, the antecedent of the metalinguistic anaphor is a quotation. This means that both the antecedent and the anaphor refer to a linguistic entity (the same one, it turns out; these expressions are co-referential). In the second subclass, the antecedent is not a quotation; it is a string in ordinary use. Here we have no co-reference: whereas the anaphor refers metalinguistically, the antecedent either refers to

*I am grateful to the editors for allowing me to submit this paper with a year's delay. I also wish to thank two anonymous reviewers for their useful suggestions and for helping me improve the design of this paper.

an object in the world or does not refer at all. This second subclass is especially interesting because it instantiates a shift in the universe of discourse, from extralinguistic reality to language. Where such a shift occurs, I will speak of 'world-to-language' anaphora.

I will argue that metalinguistic anaphora is best described in terms of a theory that assumes that various anaphoric expressions encode various degrees of salience of referents. But I will also show that salience is built in the context of utterance. It is not necessarily an acquired feature of the referent by the time the anaphor is processed: there is adjustment between the anaphor and its immediate linguistic environment. Besides, we will see that other factors may also affect anaphora resolution, which suggests that the best account must, in essence, be pragmatic.

1. Formal varieties of metalinguistic anaphora

I start with a couple of examples of 'unshifted' metalinguistic anaphora:

(4) "Harry said 'I didn't do that' but he said it in a funny way"

(5) "'You are wrong'. That's exactly what she said"

Both examples are from Levinson, who writes about (4) that "*it* does not refer to the proposition expressed but to Harry's utterance itself" (2004: 103). Note that *'I didn't do that'*, being a direct quotation, itself refers to Harry's utterance. Thus, *it* and its antecedent are co-referential and, in that respect (4) – like (5) – contains an ordinary case of anaphora.

World-to-language anaphora is less well-behaved. There is no co-reference between the antecedent and the anaphor, in which respect world-to-language anaphora belongs with bridging (cf. Clark 1977). There is, however, a major difference between bridging and metalinguistic anaphora: in all varieties of bridging that I am aware of, there are strong constraints on the linguistic form of the antecedent and of the anaphor. Thus, to mention two examples, Kleiber's (1999) 'associative anaphors' must be definite NPs with an NP-antecedent, and Erkü and Gundel's (1987) 'indirect anaphors' must be non-pronominal NPs whose antecedent quantifies over individuals, events, situations and facts. As we will see shortly, there are no such constraints on world-to-language anaphora: the antecedent merely needs to have been uttered (not too long before processing of the anaphor – this is a minimal temporal or spatial constraint) and the anaphor can, at first blush, be an indefinite NP, a definite NP, a demonstrative NP or pronoun, a relative or interrogative pronoun, and even an unstressed personal pronoun. Here are illustrations of each of these cases:

(6) They genuinely tried to become, to use **a horrid word**, acculturated with the white invaders, even if they had no desire to be assimilated. (BNC AJV 758)

(7) I still remember the day, before he was repatriated (Ray explained the meaning of **the word** to me very carefully) back to Paris by the French Government for treatment […].
(www.brain.riken.go.jp/bsi-news/bsinews18/no18/networke.html)

(8) A: I think of him as a family man.
B: Funny, I've always considered **that phrase** an oxymoron.
(Julian Barnes 1998, *England, England*, Picador, p. 64)

(9) 'Yeah, you're all right. But you're not perfect, and you're certainly not happy. So what happens if you *get* happy, and yes I know **that**'s the title of an Elvis Costello album, I used the reference deliberately […]'. (Nick Hornby 1995, *High Fidelity*, Indigo, p. 223)

(10) Yes, everything went swimmingly, **which** is a very peculiar adverb to apply to a social event, considering how most human beings swim. (Julian Barnes, *Love, etc.*, Picador, pp. 70–71)

(11) It means nothing to you, I suppose, he said, it was just a, **what** do they call it, a one-night-stand. (David Lodge, *Nice Work*, Penguin, p. 297)

(12) After several hours of bouncing from one bureaucrat (notice **it**'s a French word) to another I was allowed into the hallowed chambers.
(www.vt-fcgs.org/miscinfo.html)

In all these instances, there is a strong connection between a heteronymously mentioning expression ('heteronym', for short)[1] and some string occurring in the co-text. To that extent the strings in question can be regarded as antecedents. I wish to argue, however, that indefinite NPs are different and are not in fact anaphors. The metalinguistic NP in (6), *a horrid word*, occurs as part of an elliptical parenthetical clause. When completed, that clause is something like *I am going to use a horrid word*, and it has truth-conditions to the effect that the speaker has to use a horrid word. There is no constraint on *which* horrid word should be used: the word *acculturated* is not part of the truth-conditions of the parenthetical. In other words, the heteronym does not substitute anaphorically for its 'antecedent'. Actually, this is the conclusion one is led to every time an indefinite NP is used metalinguistically. Here are two more examples:

(13) The copper-haired woman, meanwhile, had almost canceled a Hawaii trip because of fear of terrorists (**a word** she pronounced with two syllables, like Laura Bush), […]. (starbulletin.com/2002/02/17/travel/story2.html)

1. The boldtype for the heteronyms is an addition of mine.

(14) And as the books about Peter Cook state, he was heavily influenced by the satirical nightclubs (always **an odd phrase**, to my mind) of France and Germany. (groups.yahoo.com/group/peter_cook/message/4052).

Here again the indefinite metalinguistic NPs occur in elliptical parentheticals. In the fleshed out clauses (*This is a word which she pronounced with two syllables* and *This is always an odd phrase, to my mind*), the indefinites do not refer to the strings *terrorists* and *satirical nightclubs*. This is all the clearer here because those clauses do include heteronymous anaphors to these expressions: the demonstrative pronouns that have been filled in. It is those demonstratives, not the indefinite metalinguistic NPs that contribute the strings *terrorists* and *satirical nightclubs* to the truth-conditions of the fleshed out clauses.

Things are significantly different in all of the other examples (7–12). There, the heteronym is necessarily interpreted in terms of the antecedent. For instance, what *that phrase* contributes to the truth-conditions in (8) is the expression *family man* and what *which* contributes to the truth-conditions in (10) is the word *swimmingly*. Therefore, in the rest of this paper, I will exclude examples like (6), (13) and (14) from the study of metalinguistic anaphora.

2. Constraints on the antecedent and referent of the antecedent

We have just seen that the form of metalinguistic anaphors is relatively unconstrained. But are there perhaps stronger constraints on *the sort of antecedent* that these anaphors can take, or on *the referent* of the antecedent? First, what few constraints there are on the antecedents are not very severe: the antecedent must (i) be a linguistic expression and (ii) be close to the world-to-language anaphor.[2] As for the referent of the antecedent, there are no restrictions on it … because there is not even a requirement that the antecedent should *have* a referent. Take (7) and (9) again: neither *repatriated* nor *get happy*, which are the respective antecedents of *the word* and of *that*, have a referent. This does not prevent these expressions from functioning as antecedents of the anaphors.

2. How close is an important question that I cannot answer now. Some work in psycholinguistics provides pointers, however. Levelt (1989: 122) writes that it is likely that "in conversation, literal recall not supported by salient content or pragmatic significance is short-lived, probably going back only as far as the last clause uttered". Note too that *what* in (11) precedes its 'antecedent', indicating at least that the antecedent need not always come immediately before the anaphor (which some would therefore call a *cataphor*. Still, I will follow a widely accepted tradition (see e.g. Huddleston & Pullum 2002: 1455) and use *anaphor* regardless of the position of the antecedent).

3. A top-down approach

The above observations mean that there is little sense in working out a bottom-up account based on the sorts of forms that enter into metalinguistic anaphora. It makes better sense to approach it top-down, starting from a general cognitive or pragmatic principle. There are at least two theories in the literature that attempt to ground anaphora resolution on a cognitive principle. One is Mira Ariel's 'Accessibility Theory' (1988, 1991), the other Gundel et al.'s 'Givenness Hierarchy' (1993, 2003; Borthen et al. 1997). Both share the view that different types of grammatical forms or constructions (personal or demonstrative pronouns, definite descriptions, etc.) encode different degrees of salience[3] of referents. In other words, the required degree of salience of a referent is part and parcel of the lexical meaning of referring expressions in general and anaphors in particular.

In the following, I will refer mainly to the Givenness Hierarchy, though the discussion could be extended to Ariel's Accessibility Theory.[4] The Givenness Hierarchy relies on the notion that the choice of an anaphoric form reflects the speaker's assumptions about how salient the referent of the antecedent is to the hearer (how easily recoverable it is). Gundel et al. distinguish six levels pairing a 'cognitive status' with linguistic forms. I illustrate levels 3 to 6, the other two being irrelevant to my present purposes as they are not available in metalinguistic anaphora.

As one moves up the hierarchy, one encounters forms that place increasing constraints on the cognitive status of the referent:

(15) [level 3] Steve's car let him down yesterday. **The battery** was dead.

(16) [level 4] Steve's car let him down yesterday. **That battery** was dead.

(17) [level 5] ??Steve's car let him down yesterday. **That** was dead.

(18) [level 6] ??Steve's car let him down yesterday. **It** was dead.

(15) is a basic instance of bridging. All that is required for the referent of *the battery* to be accessed is that it be 'uniquely identifiable' by the hearer. It must be recoverable as a distinct object in the context of utterance, but need not have been represented in the hearer's mind to begin with. The only condition is possession

3. Other possible terms here include *accessibility* and *activation*. However, for the sake of convenience, I will use *salience* throughout.

4. Ariel views accessibility as a property of the mental representations of referents, rather than of referents themselves. But that difference has no impact that I can see on the account given here.

of a mental frame specifying that cars have batteries. For (16) to be acceptable, we need something more: the referent must be 'familiar' to the hearer, i.e. represented in his long-term memory. The only way to interpret *that battery* here is as *that battery we've already talked about* or some such phrase. As for (17) and (18), it is very unlikely that they license the intended interpretation, i.e. that on which the anaphor refers to the battery in Steve's car. According to Gundel et al., that would be because the previous co-text is insufficient to enable the battery to be 'activated' (i.e. placed in short-term memory) at level 5 or 'in focus' at level 6 (i.e. "at the current center of attention", among the "entities which are likely to be continued as topics of subsequent utterances" (Gundel et al. 1993: 279)).

The cognitive constraints on the highest level in the hierarchy are quite severe. Thus, Gundel et al. have shown that even explicit mention does not guarantee in-focus status. Consider the next pair of examples (assume that they are uttered with the same unmarked intonation):

(19) I've just bought a parrot from Peru. **It**'s a wonderful bird.

(20) ??I've just bought a parrot from Peru. **It**'s a wonderful country.

After the first sentence in (19) has been processed, the parrot from Peru is in focus because the phrase mentioning it occurs as a direct object of *bought*, a position which, like the subject position, is capable of bringing a referent into focus. By contrast, (20) is less clearly acceptable (or requires an enriched context, in which, for example, Peru was the topic before (20) was uttered, or in which *Peru* is given prosodic prominence). Reduced acceptability stems from the fact that *Peru* occurs in a prepositional phrase modifying the head of the direct-object NP, a syntactic position that is not in itself sufficient to bring a referent into focus.

4. What endows linguistic referents with the required level of salience?

In this section, we need a strict distinction between world-to-language and unshifted metalinguistic anaphora. In the latter case, it may be assumed that the linguistic referent derives its salience from being *mentioned*. Take:

(21) The term "berber," while still used by some, is problematic. **The term** is of Greek derivation, meaning "foreigner" or "non-Greek speaker." (www.amazighworld.org/communication/who/abouttheportal.php)

The occurrence of *The term* in bold face is anaphoric: it is to be interpreted in terms of the previous occurrence of (*The term*) *"berber"*. Here anaphora resolution seems a straightforward affair since the antecedent has metalinguistic reference and already mentions the referent of the anaphor.

World-to-language anaphors are a different kettle of fish, precisely because the domain of reference shifts between the antecedent and the anaphor. In (7) to (12), the metalinguistic anaphors have antecedents that are not mentioned but *in ordinary use*. How come they are available at all for anaphora?

The question becomes all the more pressing when one realises that even unstressed personal pronouns can occur as world-to-language anaphors. As Ariel states, the speaker who uses an anaphoric personal pronoun assumes that the mental representation to be retrieved is *highly accessible* (1988: 77, 1991: 449). We saw above that Gundel et al. make a similar claim. In a recent contribution (2003: 284), Gundel, with other collaborators, goes so far as to say that "[…] unstressed personal pronouns, including *it*, are appropriately used only when the referent can be assumed to be in focus for the addressee *prior to processing of the referring form*" (italics mine).[5]

If these claims are correct, the prediction is that any unstressed personal pronoun occurring as a successfully interpretable world-to-language anaphor signals that its referent (some linguistic entity) *was already in focus*. This must be the case in sentence (12), repeated below, as well as in (22):

(12) After several hours of bouncing from one bureaucrat (notice **it**'s a French word) to another I was allowed into the hallowed chambers.

(22) This grinder uses a dead man switch to activate the grinder. **It**'s what we call it because as soon as you release pressure the switch turns off. (www.wholelattelove.com/buyingguide.cfm?buyingguideID=4)

I gather that these utterances are not especially difficult to interpret, that anaphora resolution takes place quite smoothly and that ordinary hearers/readers would not notice anything special going on. None the less, it is a fact that, in (12) and (22), *it* heteronymously refers to a linguistic entity about which nothing has been *said* in the previous context. Nor, for that matter, have the expressions *dead man switch* and *bureaucrat* been brought into special prominence, one way or another (e.g. by scare quotes in writing, by prosody in speech).

It may seem then that all it takes for *dead man switch* and *bureaucrat* to be in focus (i.e. available for reference by means of *it*) is for these words to have been uttered, i.e. to have been made perceptible in the preceding co-text. This in itself is an intriguing proposal. But there is more: the entities that are made manifest in the co-text are not the actual referents of the two occurrences of *it*: these anaphors refer to expression-*types*, not to the particular tokens occurring in an utterance

5. A very similar position is expressed in Cornish (1999: 6).

of (12) or (22). It is to these expressions as types (here, as lexical items) that the predicates *call* and *French word* apply.

To sum up, we are faced with two main questions. First, there is a question that concerns all instances of world-to-language anaphora: How come those linguistic entities are available for reference at all? This issue needs to be addressed regardless of which anaphoric expression is used. Second, we need to ask how those linguistic entities can sometimes be so highly activated that they are felicitously picked up by cognitively demanding anaphors, notably by unstressed personal pronouns. While trying to answer these questions, especially the second one, I will have to say something about how salience is built in the context of a discourse.

4.1 What enables those linguistic entities to function as referents?

A plausible answer to that question can be found in a 1998 paper on quotation by Paul Saka. There, the author puts forward a 'multiple ostension' hypothesis according to which any expression used in a spoken utterance 'directly ostends' a phonetic token (say /dʒɒn/) and 'deferringly ostends' a number of other features or entities: when I utter the word *John*, I explicitly refer to John (that is part of what I say), but I also 'activate' (via the phonetic token) the corresponding form type, the lexeme ⟨*John*, /dʒɒn/, proper noun, JOHN⟩, the concept JOHN (see Saka 1998: 126).

I think Saka's hypothesis is sound. Therefore, I assume that the result of uttering any expression is that various linguistic aspects of it become objects of the context of utterance (of the universe of discourse) endowed with at least a minimal degree of salience. These objects then are potential targets for subsequent referential expressions because the hearer/reader is at least minimally alert to them as the discourse unfolds.

4.2 What brings linguistic entities into focus?

As far as I can see, in world-to-language anaphora, the linguistic referent (an expression-type, a lexeme, a meaning) possesses no other salience than that which it acquires indirectly from the uttering of an associated expression-token. That is all the salience it has *prior* to the moment when a subsequent world-to-language anaphor is processed.

This, at least, holds for cases where no specific intonation pattern or typographical markers have been used. I return to these highlighting devices shortly, but for the moment I rest satisfied with the claim that the antecedents in (12) and (22) can all be uttered neutrally and still be recoverable as antecedents of metalinguistic anaphors. I assume that most of my readers had no trouble arriving at the correct interpretation for these utterances, even though the antecedents were not highlighted in any way.

Now, if Saka's story accounts for all the salience of the relevant linguistic referents, then it must also singlehandedly account for the *level* of activation of these referents. Thus, with respect to cases like (12) and (22), where in-focus status is required, the theory in Gundel et al. (2003) predicts that multiple ostension (or some similar story) adequately explains how that demanding cognitive status is achieved here. Yet, this prediction cannot be right, for several reasons. First, we saw earlier (Examples 19, 20) that the explicit mention of a referent may not suffice to bring it into focus. One has reason to doubt that mere deferred ostension will succeed where even explicit mention may fail. Second, it is not sensible to assume that the exact same event (some words have been uttered) can lead to such different outcomes as the following: making a referent 'uniquely identifiable' (when the anaphor is a definite NP), placing it in short-term memory (demonstrative pronouns), or bringing it into focus (unstressed pronouns). Third, and probably worst of all, Gundel et al.'s prediction would seem to have an absurd consequence: since all (the entities associated with) the tokens uttered prior to the anaphor in (12) or (22) are ostended to the same extent (as a direct result of being uttered), this must mean that they are *all* in focus (for at least some period of time). Clearly, this is an undesirable conclusion.[6]

So Saka's story cannot be all there is to the salience of *bureaucrat* in (12) and *dead man switch* in (22). But since the multiple ostension hypothesis is all we have to explain what happens *to the left* of the anaphor, we must look for an explanation *to its right*. This explanation is disappointingly straightforward: it lies in the combination of the anaphor with a metalinguistic predicate. Try replacing that predicate with a neutral one that applies equally well to linguistic and non-linguistic objects, and the anaphor can now hardly be interpreted as a case of heteronymous mention:

(23) I was attacked by Zonkins.

a. **It**'s a strange word.
b. How do you spell **it**?
c. **It**'s strange, isn't it?
d. How do you like **it**?

(24) This grinder uses a dead man switch to activate the grinder. **It**'s funny, isn't it?

(25) After several hours of bouncing from one bureaucrat (notice **it**'s French) to another I was allowed into the hallowed chambers.

---

6. As a reviewer notes, the three reasons given may be no more than different ways of identifying a single cause, namely, the fact that the degree of salience of a referent can still change *after* the occurrence of the expression that makes that referent available. See below.

Whereas, in (23a–b), *it* is naturally construed as referring to the name *Zonkins*, it can hardly be understood to do so in (23c–d). Yet, the predicates *strange* and *like* do not bar the metalinguistic interpretation, since they are neutral with respect to a world-oriented or language-oriented reading. In (24), *it* is now very unlikely to be interpreted as referring to the phrase *dead man switch*. (25), it appears, is different. The preferred interpretation for *it* seems to be heteronymous mention. Some members of the audience at the 2005 Dortmund conference on Constraints in Discourse suggested that *French* might be more 'language-biased' than the other alleged neutral predicates. That may be so. But note that, in the absence of additional contextual information, an utterance of *Champagne is French* is ambiguous between a world-oriented and a language-oriented reading (with a preference for the former). This casts doubt on that explanation. My feeling is rather that the heteronymous reading is facilitated by the occurrence of *it* in the parenthetical clause *notice it's French*, which intrudes at an arbitrary spot in the sentence structure: in this case, it splits a place adjunct into two fragments, thereby disrupting the humdrum flow of information and introducing a new topic. This intrusion may draw attention to the words to its immediate left, one of which is indeed French, increasing the likelihood that *it* is going to be rightly interpreted as shifting reference from the world to language.[7]

The facts about metalinguistic predicates in (23)–(25) suggest that *givenness*, in the strict sense, is inadequate to explain world-to-language anaphors. The same applies to accessibility understood strictly as *prior accessibility*. Most likely, in (23)–(25), the referent is not in focus *before* the occurrence of anaphoric *it*. Here, therefore, is my preferred account of what goes on:

> the anaphor, being an anaphor, triggers a presupposition that its referent has already been introduced into the discourse. The presupposition initiates a search for this referent. Processing of the metalinguistic predicate then helps direct the addressee's attention to linguistic entities. That way, a linguistic referent is *brought into focus*.

5. Role of the predicate elsewhere

The previous discussion has shown that givenness (or accessibility) cannot in and of itself explain how pronominal metalinguistic anaphors select their referent. I have suggested that, in order to account for the pronoun's ability to pick out a

7. At this stage, these are only sketchy and speculative remarks that need substantiating. But a proper study of the role of parentheticals would take me too far afield.

linguistic referent, it is necessary to look to the right and consider the metalinguistic predicate. My main aim has been to show the inadequacy of givenness/accessibility as a unique explanatory factor. In Section 7, I outline a general framework for anaphora resolution that takes account of the contributions of givenness/accessibility and of the predicate with which the anaphor combines, but also allows for other factors. In the meantime, however, I wish to look further into the role of the metalinguistic predicate.

It is striking that neither of the theories I have appealed to has much to say about the lexical contribution of the predicate governing the anaphor, or, more generally, about the contribution of the co-text to the right of the anaphor. However, they might be justified in neglecting the right-hand co-text provided, simply, that the predicate barely played a role at all in the sorts of anaphora that they were considering (i.e. outside world-to-language anaphora).

But that is just not true. In a study of referents introduced by clauses rather than NPs, Gundel et al. do treat the predicate as a determining factor. Still, they do so only implicitly:

(26) a. John insulted the ambassador. **It** happened at noon.
b. John insulted the ambassador. ??**It** was intolerable. (cf. Gundel et al. 2003: 285)

The claim here is that some clausally introduced entities license subsequent reference by means of *it* while others do not. Thus, if the entity in question is an event, *it* should be usable, whereas if the entity is a situation (or 'fact', in some terminologies), *it* should be dispreferred and, for instance, replaced by *that*. Variation in the acceptability of anaphoric *it* means that described events are more highly activated than described situations. In the case at hand, if *John insulted the ambassador* introduces a situation rather than an event, then it is more difficult to refer to this situation by means of the personal pronoun *it*.

The problem is that, by the time *John insulted the ambassador* has been processed, it is still an open question whether that sentence denotes an event or a situation. This, by the same token, also means that the degree of salience of the described entity is not fixed yet. Not until the predicate of the following sentence has been processed can the first sentence be said to denote an event or a situation. Thus, (26a) is construed as an event because of the subsequent occurrence of *happened*. Similarly, interpretation of *intolerable* in (26b) turns John's insulting the ambassador into a situation. In 'salience-talk', we are induced to say that the degree of salience of the entity introduced by *John insulted the ambassador* cannot be determined until *happened* or *intolerable* have been processed. In particular, it appears that the in-focus status of the event of John's insulting the ambassador depends in part on the occurrence of the verb *happened* in the right-hand co-text.

These examples show that the influence of predicates on the salience of their anaphoric arguments extends beyond world-to-language anaphora. But, as I said, it is striking that Gundel et. al. though their analyses implicitly acknowledge this, do not point out the role of the predicates to the right of anaphors. Neither, for that matter, does Ariel in the papers mentioned earlier.

6. Robustness of the proposed account

One way of testing the role played by the metalinguistic predicate is to check what happens when the salience of the linguistic referents is enhanced by means other than a metalinguistic predicate.

I hinted above that antecedents could be highlighted in such a way as to attract hearers' attention to them *qua* expressions. However, the following set of examples suggests that such highlighting is not enough to endow a linguistic referent with the required degree of salience:

(27) I was attacked by 'Zonkins'.
  a. **It**'s strange, isn't it?
  b. How do you like **it?**

(28) This grinder uses a *dead man switch* to activate the grinder. **It**'s funny, isn't it?

Even if the scare quotes or italics are realised prosodically as very emphatic stress, it is unclear that construal of *it* as a heteronym is facilitated.

Some further evidence for the role of the metalinguistic predicate comes from 'the other half' of metalinguistic anaphora, i.e. the 'unshifted' cases. Consider what happens to (29) if the metalinguistic predicate is replaced with a neutral one:

(29) "We're in a forest, with spiders and who knows what other yuckies…" **this** she said as she wrinkled her nose. (www.ofelvesandmen.com/Stories/D/Di-and-LK/RoadTrip2–3.htm)

(30) "We're in a forest, with spiders and who knows what other yuckies…"
  a. **this** she hated, so she wrinkled her nose.
  b. **that** was pretty bad.[8]

In (29) and (30), the linguistic referent – the female character's utterance – is mentioned by means of the direct quotation. Yet, once again, it appears that without a metalinguistic predicate the heteronymous interpretation of the anaphor is

8. The predicates *hate* and *be pretty bad* are neutral in the sense that they can apply to linguistic entities just as well as they do to extralinguistic ones. (cf. the predicates used in (23)–(25)).

hardly available. Notice moreover that we are dealing with demonstrative pronouns, i.e. *less* cognitively demanding forms than personal pronouns, requiring only that the referent be activated (30a) or familiar (30b), according to Gundel et al. (1993). Yet, even though the referent is mentioned explicitly, and the anaphoric forms demand less salience, the intended interpretation cannot be readily accessed.

Things, however, are not as clear-cut as with world-to-language anaphors. Take Example (21) again and consider what happens if a pronoun is substituted for the definite NP and a neutral predicate for the metalinguistic one:

(21) The term "berber," while still used by some, is problematic. **The term** is of Greek derivation, meaning "foreigner" or "non-Greek speaker."

(31) The term "berber," while still used by some, is problematic. **It** is strange/We don't like **it**.

Here, it seems that something like default assumptions about topic continuity ensure that *it* is going to be interpreted correctly as a heteronym. Alternatively, the presence of the introductory predicate *the term* in the antecedent sentence may affect the availability of the heteronymous interpretation of *it* in (31). Note also that if *this* or *that* were used instead of *it*, the metalinguistic interpretation would not be secured. That can probably be explained by (something like) Grice's maxim of quantity: if a more cognitively demanding form can be used, then a less demanding one cannot be used without a special reason: a plausible candidate here is topic change. Clearly, further work needs to be done on the factors that affect the availability of a metalinguistic interpretation for an unshifted anaphor. But this goes beyond the scope of the present study.

7. Plausible explanations

In this section, I will extend the argument that prior accessibility or salience cannot be the sole explanatory factor for metalinguistic anaphora resolution. In 7.1 and 7.2 I discuss two more factors that may play a part, the lexical meaning of the antecedent, and the possibility that some linguistic referents are inherently more salient than others. The conclusion I am led to is that all these factors need to be brought together into a general framework, something along the lines of Recanati (2004). This is not incompatible with Gundel et al. and Ariel. Instead, it is to be seen as an extension of their theories, one that further develops the notion that anaphora resolution is primarily a pragmatic affair.

The story so far offers an explanation for why *a* referent can be made salient enough. But it does not say how a *particular*, non-random referent is picked out.

Access to the referent is provided via identification of the antecedent, but how is this antecedent selected?

7.1 The lexical meaning of the antecedent

First, there are instances in which the lexical meanings of both the antecedent and the metalinguistic predicate play a major part. Take (8) again:

(8) A: I think of him as a family man.
B: Funny, I've always considered **that phrase** an oxymoron.

Since *an oxymoron* is predicated of *that phrase*, the latter NP must refer to something that can be judged to be both a phrase and an oxymoron. Therefore, the antecedent must (i) be a complex phrase and (ii) denote two properties that can, in a given context, be interpreted as incompatible. Only an antecedent whose lexical meaning is consistent with that of the predicate will be selected. If there is no more than one such antecedent – here, there is only *family man* – then anaphora resolution is straightforward. In any case, all words or phrases whose meaning did not match that of the predicate would already have been eliminated.

However, the match between lexical meanings does not systematically play a part in anaphora resolution. Thus, in (9), this explanation is entirely unavailable, because the metalinguistic predicate, *title of an Elvis Costello album*, does not favour any particular antecedent: only a fully context-based, pragmatic, explanation can do the job. The fact that the antecedent of *that* is *get happy*, rather than the italicised *get* or *happy* is entirely a matter of world knowledge. Anyone unfamiliar with the Elvis Costello catalogue would be unable to pick out the right referent.

7.2 Inherent salience of certain words or phrases

In discussion, there have been suggestions that some antecedents inherently stand out. Take rare words, like *swimmingly* in (10), or long ones like *repatriated* in (7). I do not deny that rare or unusually long words tend to attract the addressee's attention. Some psycholinguists at least have argued that different words have different activation thresholds (cf. 'the frequency effect'; see Garman 1991: 279–81 or Harley 1995: 71–73). All the same, words are not *inherently* salient, and the 'lexical oddity' thesis therefore falls short of being a satisfactory explanation. High salience is not part and parcel of a lexeme, it is context-sensitive. Two quick examples: specialised, technical words tend to be rare words. This means that they are likely to attract attention when used in an everyday conversational context. Yet, when used in a specialised one, these terms will not be particularly salient. Or consider the fact that even the most ordinary words can become highly salient, provided they are uttered in the right context. Thus the word *baby* heard by a man who is about to become a father.

7.3 A general pragmatic framework

The factors that have so far been shown to play *some* part in the resolution of metalinguistic anaphora are:

- the accessibility of the linguistic referents
- a constraint on the distance between antecedent and anaphor (end of Section 2)
- the match between the lexical meaning of the antecedent and the predicate with which the anaphor combines (or which is included in the anaphor, in the case of definite and demonstrative NPs)
- unequal threshold levels of contextual salience

All these factors are potentially relevant, and they need to be integrated into a general framework. This framework must be essentially pragmatic, as it appears that no single parameter can account for metalinguistic anaphora resolution: although some factors (distance, for instance) probably play a role in all cases of resolution, it is usually a *cluster of factors* that combine to help identify the right antecedent and select the corresponding referent. And which factors are relevant appears to be context-dependent.

Let us once again focus on pronominal world-to-language anaphors. What happens when a hearer/reader comes across the anaphors in utterances like (9)–(12) or (22)? It is very likely that the hearer/reader will not immediately ascribe to these a linguistic referent. On the contrary, he is likely at first to have in mind an ordinary extralinguistic referent. That is because nothing in the co-text *up to that point* has particularly activated any linguistic entities. Thus, in (9), he would probably expect *that* to refer to the event of getting happy (cf. *So what happens if you* get *happy, and yes I know that's not likely to happen soon*). Similarly, in (10), the default interpretation for *which* is likely to be the fact that everything went swimmingly (cf. *Yes, everything went swimmingly, which was rather a surprise*).[9]

At this stage, two stories are, I think, conceivable. On the first, a costly 'repair' procedure is assumed to take place: the pronominal anaphor's default interpretation enters into semantic composition with the interpretation of the metalinguistic predicate. This yields an absurd interpretation for the whole clause (something to the effect that "the event of getting happy is the title of an Elvis Costello album" or that "the fact that everything went swimmingly is a peculiar adverb"). The absurdity of this result then induces the hearer/reader to backtrack and re-interpret the

9. That there has been no particular previous activation of a linguistic-entity-as-referent is patently true, of course, when the anaphor *precedes* its antecedent, as in Example (11).

pronoun as having metalinguistic reference. Only on this second attempt does the hearer/reader come up with an acceptable interpretation for the whole clause.

The second account does not assume such a cognitively costly procedure. Here, the idea is that the extralinguistic interpretation of the anaphor does not undergo composition: it is 'entertained' only until the hearer/reader encounters the metalinguistic predicate, at which stage it is superseded by the metalinguistic reading. The assumption is that this reading was already activated 'by association': as soon as words are uttered, they (or their related types) are endowed with a minimal degree of activation – this is the gist of Saka's hypothesis. Therefore, what happens is that the linguistic referent enters directly into semantic composition with the metalinguistic predicate: there is no cancellation of an initial, primary interpretation *of the whole clause*. The first time the whole clause is interpreted, it is *qua* referring to a linguistic entity.[10]

The second account is essentially that given for figurative language by François Recanati in his *Literal Meaning* (2004: 27–30). Recanati rejects a dominant 'Gricean' model according to which an absurd literal interpretation for the whole utterance is processed first, then discarded because it fails to comply with some conversational maxim, thus triggering a repair procedure by means of conversational implicatures. Recanati outlines a different story, one that makes allowances for 'primary pragmatic processes', i.e. associative processes by which the meaning of local (subclausal) constituents is adjusted *before* undergoing semantic composition.

Interestingly, Recanati then shows how this model can be extended to the selection of a referent for an anaphoric pronoun (2004: 31–32) and explicitly underlines the role that processing of the predicate plays in affecting the accessibility of the candidate referents (2004: 31, 33). My account of world-to-language anaphora is in the same spirit. In particular, I take it that the processes involved in interpreting pronominal anaphors are cognitively the same as those described by Recanati: they are associative rather than properly inferential; they apply locally rather than to the whole utterance. However, the analogy is perhaps only partial. The processes described by Recanati are not 'pre-semantic'; they are not part of the processes that lead to determining 'which sentence was uttered'. By contrast, it is reasonable to hold that choosing between an ordinary and a metalinguistic reading of a pronoun may be a facet of disambiguation. Assume I say *I love Chicago!*. I may

10. This paragraph is a follow-up on a comment made by Rachel Giora (p.c.). Giora suggested that my account would predict garden-path effects in cases of world-to-language anaphora, adding that such effects could be tested empirically. I have not had the opportunity to do this, but I think that my account leaves open the possibility that world-to-language anaphors are not garden-path sentences, that some adjustment of anaphor and predicate takes place 'locally' before the whole sentence is processed.

mean the proper name to refer to a variety of objects, notably a city, a word (e.g. if I like the pronunciation), but also a musical, a band, etc. It seems to me at least that the sentence about the city is a different sentence from the one that refers to the word. If that is correct, selection between an object-level and a meta-level referent for *Chicago* is part of disambiguation. And if, as seems sensible, one extends this analysis to all cases in which a linguistic referent competes with an extralinguistic one, there is a genuine difference between my account and Recanati's.

8. Binding or accommodation?

I believe the story told so far can be reformulated in terms of dynamic models of discourse such as DRT. The cognitive constraints associated with the use of anaphors can be cashed out in terms of presupposition. On this view, use of anaphoric *this*, for example, triggers the presupposition that, say, the referent must be present in short-term memory. And, even if it turned out that metalinguistic anaphors are really demonstrative deictics, this could be captured in presuppositional terms as well.[11] Use of indexical *this* would trigger the presupposition that the speaker is demonstrating (e.g. pointing at) the relevant referent. On theories like those set out in Geurts (1998) and Beaver & Zeevat (2007) the presuppositions triggered by anaphors and demonstratives[12] can either be 'bound' to an antecedent (by which the relevant referent has been previously introduced into the discourse and is therefore part of the 'common ground'), or 'accommodated' (see below). If neither happens, presupposition failure occurs and communication is unsuccessful.

In my examples, there can be no talk of presupposition failure: world-to-language shifts do not usually cause breakdowns in communication. Therefore, the presuppositions triggered by shifters must be accounted for in terms of

---

11. I have talked of the expressions shifting reference from the world to language as *anaphors*, thus adopting John Ross's terminology in a 1970 squib about *metalinguistic anaphora*, the first discussion of the topic known to me. Many writers, however, talk of *discourse* or *textual deixis* instead (e.g. Lyons 1977: 667; Webber 1988: 116; Huddleston & Pullum 2002: 1460–61; Levinson 2004: 103, 108). I readily admit that there is a strong kinship between my data and deixis, especially the use of demonstrative NPs. Still, evidence seems to go both ways. (For some discussion, see De Brabanter 2004.) I cannot go into this issue here, but I rest content with the idea that there may be more similarities than differences between the two phenomena, a view for which Recanati (2005) makes a convincing case.

12. In the rest of this section, I shall use the term *shifter* as a cover term (for expressions that host a shift in the universe of discourse), so as to leave the door open to a deixis-based account too.

binding or accommodation. Binding, however, is to be ruled out: the linguistic referent introduced by a shifter is new: it is not part of the common ground by the time the shifter is processed. Therefore, the prediction is that shifter-triggered presuppositions are accommodated. Accommodation comes in when a presupposition that cannot be bound is nevertheless interpretable by the hearer. Let us take a simple example. A stranger comes up to me in the street and says:

(32) I've lost my dog! Can you help me?

Use of *my* triggers the presupposition that the man has a dog. But this cannot be bound to any information already in the common ground: the man is a perfect stranger to me. Yet, I have no trouble *adding* this information to the common ground. I 'accommodate' it, and no communication failure ensues. Now consider this other example: I hear a tremendously loud bang. A stranger turns to me and exclaims:

(33) That sounded like an explosion.

Her utterance is not accompanied by any demonstration. Although Beaver & Zeevat suggest that "the use of the demonstrative presupposes the demonstration by the speaker. If no such demonstration occurs, infelicity results: the reader cannot simply incorporate a referent" (2007: 535), I none the less assume that they would agree that accommodation still takes place provided a referent is salient enough in the situation of utterance. That is the case here given the loudness of the bang.[13]

It is not clear that a similar story can be told about the presuppositions triggered by shifters. As pointed out in 4.2, there is no notable event foregrounding the intended linguistic referent, nothing like the loud bang in (33). This means that the Beaver/Zeevat account predicts presupposition failure in most cases of world-to-language shift, a prediction that cannot be correct, because hearers generally have no trouble identifying the right referent.

At this stage, the only way for a DRT-based account to address this issue is to introduce an extra layer of representation for speech events themselves,[14] with special constraints designed to capture the short span of time during which uttered linguistic forms remain accessible. A step in this direction is taken by Corblin & Laborde (2001), in their study of the French counterparts of *the former* and *the latter*.

---

13. Ginzburg 2001: 21 has a similar example (his (24a)):

[Context: a shot is heard, followed by a woman's scream:] A: Oh boy, she sounds scared.

Ginzburg's comment is that "deixis makes anaphors felicitous without an overt antecedent" (2001: 21). But this is deixis *without a demonstration*.

14. Note that the difficulties highlighted in this section do not crop up with unshifted anaphors: these can be directly bound to an antecedent that refers to linguistic material.

Another such attempt, though not within DRT, is Ginzburg & Cooper's (2004) HPSG formalisation of 'clarification ellipsis'.[15] Ginzburg & Cooper focus on cases like *A: Did Bo finagle a raise? B: finagle?*, where B's elliptical question is susceptible of either a 'clausal reading' (= *Are you asking if Bo finagled a raise (of all actions)?*) or a 'constituent reading' (= *What does it mean to finagle?*). They show that any model capable of accounting for examples of this kind must include several ingredients, two of which I will single out as they are directly relevant to an account of world-to-language anaphora. One ingredient is 'utterance reference'. The idea is that clarification ellipsis – especially in its 'constituent reading' (which is often metalinguistic) – usually involves reference to utterance *tokens* (or 'speech events'), so that the syntactic-semantic representation of clarification ellipsis "must include references to (previously occurring) utterance events" (2004: 306). Note that this is exactly what is needed in the case of world-to-language anaphors. The second constraint is termed 'sub-utterance accessibility', and it originates in the observation that "any semantically meaningful sub-utterance can be clarified using [clarification ellipsis] under conditions of phonological or partial syntactic parallelism" (2004: 306). Once again, note that a similar constraint weighs upon any model of world-to-language anaphora: the antecedent of a shifted anaphor can be just any word or phrase in the close co-text. As a matter of fact, the constraint is stronger in the case of world-to-language anaphora, because the antecedent need not even be a "semantically meaningful sub-utterance" (cf. Examples (7) and (9), where the antecedent is not a referential expression). The provision of a formalised account of metalinguistic anaphora that includes utterance reference and sub-utterance accessibility is, at this stage of my research, no more than a promissory note.

9. Conclusion

The present study is in line with the earlier proposals of Gundel et al. (1993) and Ariel (1988, 1991) in that it corroborates the idea that a cognitive principle underlies the choice of anaphoric expressions. However, though an important parameter, prior accessibility is not enough to explain what goes on in metalinguistic anaphora. Elsewhere, Wilson & Matsui (1998) have also shown that it is inadequate to account for other types of anaphora resolution.

In the end, the main dividend of this study might be this: the very unobtrusiveness of world-to-language anaphors – the fact that their resolution does not pose special problems – throws light on what participants keep track of as

15. I am grateful to an anonymous referee for providing this reference.

a discourse (e.g. a conversation) proceeds. Psycholinguists have shown that discourse participants keep some record of which words were uttered.[16] But here, we have *direct linguistic* rather than experimental psychological evidence of that assumption. Furthermore, speakers must to some extent be aware of what can be assumed to be in the minds of other participants. Otherwise, we should expect anaphora resolution to be a haphazard process. But it is not. The fact that it is often successful suggests that speakers assume other participants to temporarily store *uttered linguistic forms* over and above extralinguistic referents. It is likely that participants in a discourse temporarily open mental files for linguistic entities too, for otherwise these could not serve as anchors for anaphoric expressions.

Although this idea may appear quite commonsensical, it clashes with some views found in the literature. Thus, in a discussion of anaphoric pronouns, Mark Sainsbury writes:

> In the process of interpretation, we expect understanders to carry previous interpretations forward and for them thus to be available for solving new problems of interpretation. *It would be quite another thing to expect them to carry forward memories of precise linguistic forms*. It is a familiar phenomenon that among people who use two languages interchangeably in their conversations, it often happens that one remembers what the other said, but not in which language she said it. *Interpretation is remembered, but not linguistic form*. […] In interpreting the first part of, say, (1) ("A mosquito is buzzing about our room") we come to know that what the speaker has said is true iff there is at least one mosquito which is buzzing around our room, and *we throw away all other information about the utterance, including the words in which it was couched*. (2002: 57; emphasis mine)

As far as I can see, the present study invalidates views of this sort. Sainsbury is certainly right in assuming that propositional contents are usually remembered much longer than the linguistic material used to convey them. But he fails to allow for the difference between effects on long-term and on short-term memory. What Saka calls *multiple ostension* has a short-lived impact. The various linguistic aspects associated to an uttered token are only kept in mind for a while; long enough, however, to enable the hearer to interpret successfully a subsequent shifted anaphor, provided this anaphor occurs while the ostended items are still stored in short-term memory.

---

16. See e.g. Levelt & Kelter (1982) or Levelt (1989). The idea is that, when planning a new utterance, speakers rely on the information stored in their 'discourse record' (Levelt 1989: 111) in order to make their new contribution coherent with previous moves. Levelt & Kelter adduce experimental evidence that lexical information is among those aspects of a discourse that are stored in short-term memory.

References

Ariel, M. (1988). "Referring and accessibility". *Journal of Linguistics*, 24, 65–87.

Ariel, M. (1991). "The function of accessibility in a theory of grammar". *Journal of Pragmatics*, 16, 443–463.

Beaver, D. & Zeevat, H. (2007). "Accommodation". In Ramchand, G. & C. Reiss (eds), *Oxford Handbook of Linguistic Interfaces*. Oxford University Press, pp. 503–539.

Borthen, K., Fretheim, T. & Gundel, J. (1997). "What brings a higher-order entity into focus of attention? Sentential pronouns in English and Norwegian". In Mitkov, R. & B. Boguraev (eds), *Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts*. Madrid, pp. 88–93.

Clark, H.H. (1977). "Bridging". In Johnson-Laird, P.N. & P.C. Watson (eds), *Thinking: Readings in Cognitive Science*. Cambridge, Cambridge University Press, pp. 411–420.

Corblin, F. & Laborde, M-Ch. (2001). "Anaphore nominale et référence mentionnelle: *le premier, le second, l'un et l'autre*". In W. De Mulder et al. (eds), *Anaphores pronominales et nominales. Etudes pragma-sémantiques*. Rodopi, 99–121.

Cornish, F. (1999). *Anaphora, Discourse, and Understanding. Evidence from English and French*. Oxford, Clarendon Press.

De Brabanter, P (2004). "'World-to-language' shifts between an antecedent and its pro-form". In Branco, A., McEnery, T. & R. Mitkov (eds.), *Proceedings of the 5th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 2004)*. Lisbon, Edições Colibri, pp. 51–55.

Erkü, F. & Gundel, J. (1987). "The pragmatics of indirect anaphors". In Verschueren, J. & M. Bertuccelli-Papi (eds), *The Pragmatic Perspective: Selected from the 1985 International Pragmatics Conference*. Amsterdam. John Benjamins, pp. 533–545.

Garman, M. (1990). *Psycholinguistics*. Cambridge University Press.

Geurts, B. (1998). "Presuppositions and anaphors in attitude contexts". *Linguistics and Philosophy*, 21, 545–601.

Ginzburg, J. (2001). "Clarification ellipsis and nominal anaphora". In Bunt, H., Muskens, R. & E. Thijsse (eds.), *Computing Meaning: Volume 2*. N° 77 in Studies in Linguistics and Philosophy, Dordrecht, Kluwer.

Ginzburg, J. & Cooper, R. (2004). "Clarification, ellipsis, and the nature of contextual updates in dialogue". *Linguistics and Philosophy*, 27, 297–365.

Gundel, J., Hedberg, N. & Zacharski, R. (1993). "Cognitive status and the form of referring expressions in discourse". *Language*, 69, 274–307.

Gundel, J., Hegarty, M. & Borthen, K. (2003). "Cognitive status, information structure, and pronominal reference to clausally introduced entities". *Journal of Logic, Language and Information*, 12, 281–299.

Harley, T.A. (1995). *The Psychology of Language. From Data to Theory*. Hove, Erlbaum (UK) Taylor & Francis.

Huddleston, R. & Pullum, G.K. (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press.

Kleiber, G. (1999). "Associative anaphora and part–whole relationship: The condition of alienation and the principle of ontological congruence". *Journal of Pragmatics*, 31, 339–362.

Levelt , W.J.M. (1989). *Speaking: From Intention to Articulation*. Cambridge, Mass., MIT Press, Bradford Books.

Levelt, W.J.M. & Kelter, S. (1982). "Surface form and memory in question answering". *Cognitive Psychology*, 14, 78–106.

Levinson, S. (2004). "Deixis". In Horn, L. & G. Ward (eds.), *The Handbook of Pragmatics*. Oxford, Blackwell, pp. 97–121.
Lyons, J. (1977). *Semantics*, 2 vol. Cambridge, etc., Cambridge University Press.
Quine, W.V.O (1940). *Mathematical Logic*, Cambridge, Mass., Harvard University Press.
Recanati, F. (2000). *Oratio Obliqua, Oratio Recta: An Essay on Metarepresentation*. Cambridge, Mass., MIT Press, Bradford Books.
Recanati, F. (2004). *Literal Meaning*. Cambridge, Cambridge University Press.
Recanati, F. (2005). "Deixis and anaphora". In Szabo, Z. (ed), *Semantics vs. Pragmatics*. Oxford, Oxford University Press, pp. 286–316.
Ross, J.R. (1970). "Metalinguistic Anaphora". *Linguistic Inquiry*, 1, 273.
Sainsbury, R.M. (2002). "Reference and anaphora". *Philosophical Perspectives* 16, 43–71.
Saka, P. (1998). "Quotation and the use-mention distinction". *Mind*, 107, 113–135.
Webber, B.L. (1988). "Discourse deixis: reference to discourse segments". *Proceedings of 26th Annual Meeting of the Association for Computational Linguistics*, pp. 113–122.
Wilson, D. & Matsui, T. (1998). "Recent approaches to bridging: Truth, coherence and relevance". *UCL Working Papers in Linguistics*, 10, 173–200.

Appositive Relative Clauses and their prosodic realization in spoken discourse

A corpus study of phonetic aspects in British English

Cyril Auran & Rudy Loock

Recent research on discourse has shown that Appositive Relative Clauses (ARCs) can be defined positively in spite of a long tradition in which they are defined asymmetrically with respect to Determinative Relative Clauses (DRCs). Particularly, Loock (2007) has shown that ARCs fulfill specific discourse functions, and distinguishes three categories: Relevance, Subjectivity and Continuative ARCs. This paper aims to show that, for the same syntactic structure, different functions in discourse correspond not only to specific morphosyntactic and semantic criteria, but also to different prosodic realisations.

Using attested examples taken from electronic corpora and analysed using semi-automatic procedures within Praat, this paper suggests that a distinction can be made between ARC types such as defined in Loock's study, with for instance higher onset values for subjectivity ARCs, a cue of stronger discourse discontinuity. This paper also addresses the prosodic realization of ARCs as opposed to the general category of parentheticals, which generally include ARCs.

0. Introduction

0.1 Outline

Our global project aims to relate discourse structure and functions on the one hand and prosody on the other hand. In the present study, we investigate the prosodic characteristics related to different discourse functions fulfilled by one specific syntactic structure, *viz* the appositive relative clause (henceforth ARC) in English. Using a corpus of spoken data, this work more particularly aims to show that differences in pragmatic functions correspond not only to differences in morphosyntactic and semantic characteristics but also to phonetic differences in prosodic features mainly related to intonation, rhythm and intensity, all of which are semi-automatically extracted from the speech signal. After outlining the prosodic characteristics of ARCs as a whole – and therefore shedding light on their atypical status –, we will investigate the link between the different discourse functions fulfilled by ARCs and their prosodic realisations.

0.2 Methodology

On a global scale, the research project which constitutes the framework of this particular pilot study relies on two spoken British English corpora: Aix-MARSEC (cf. Auran, Bouzon & Hirst (2004)) and ICE-GB (cf. Greenbaum (1996)). This allows extensive access to a wide variety of speech types and levels of spontaneousness. Due to availability reasons, the ICE-GB could not be used in the present study, which is thus restricted to the more formal and scripted speech types to be found in the Aix-MARSEC. This in turn induces a bias in both the representativeness and the distribution of ARCs itself.

The Aix-MARSEC constitutes a second evolution from the original SEC (*Spoken English Corpus*, cf. Knowles, Wichmann & Alderson (1996)), the MARSEC (*Machine Readable SEC*) constituting the first one (Roach et al. (1993)). The data represent more than five and a half hours of natural-sounding British English (BBC recordings from the 1980s) from 53 different speakers. The corpus contains about 55.000 orthographically transcribed and manually aligned words, manual prosodic annotation of all the recordings (using tonetic stress marks) and CLAWS I tagging and parsing of the data. Automatic procedures were used within the Aix-MARSEC project to transcribe the 55.000 words of the corpus into phonemes (SAMPA and IPA alphabets), to optimize and align this transcription with the speech signal and to group and code phonemes into sub-syllabic constituents (onset, nucleus and coda), syllabic units, rhythmic groups and intonation units. The coding of intonation was carried out using the MOMEL-INTSINT methodology developed in Aix-en-Provence (cf. Hirst et al. (2000)).

This paper more specifically focuses on the prosodic marking of elements within Loock's (2005, 2007) taxonomy of appositive relative clauses, based on differences in discourse functions and illustrated with morphosyntactic and semantic criteria. Unpunctuated written transcriptions of recordings from the Aix-MARSEC corpus (cf. Auran, Bouzon & Hirst (2004)) were manually annotated, thus leading to the identification of a sample of ARCs; we shall call this part of the procedure *discourse annotation*. The second part of the procedure, which we shall call *prosodic annotation*, then consisted in semi-automatically analysing the corresponding recordings using original scripts within Praat (cf. Boersma (2001); Boersma & Weenink (2006)). Both graphical and formal statistical analyses were eventually carried out within the R environment and software.

Due to the above-mentioned limitations, this paper will present preliminary results and tendencies concerning prosodic characteristics of two types of ARCs within Loock's taxonomy (namely *relevance* and *subjectivity* ARCs), which we now turn to.

1. Appositive Relative Clauses and their functions in discourse

In Loock (2003, 2005, 2007), we have shown that ARCs (also called non-restrictive, see 1a) can be defined positively in spite of a long tradition of asymmetrical definition with Determinative Relative Clauses (also called restrictive, henceforth DRCs, see 1b), according to which ARCs fulfil the functions that DRCs do not, hence labels like non-restrictive and non-defining.

(1) a. *The people of Oz, who were scared of the Witch of the East, were relieved when Dorothy's porch crushed her to death.* (ARC)
b. *The people of Oz who were scared of the Witch of the East were relieved when Dorothy's porch crushed her to death.* (DRC)

ARCs fulfil specific discourse functions which can be defined regardless of the long-established ARC-DRC dichotomy. Our research suggests that three main categories of ARCs can be distinguished, and that the use by speakers of an ARC, and not another syntactic structure, responds to specific constraints linked to the status of the information conveyed, in particular its discourse new/old and its hearer new/old nature, following Prince's (1981, 1992) typology of given/new information and also the previous and following context (e.g. presence of a presupposed open proposition, as defined by Prince 1986). The three main categories were labelled relevance, subjectivity and continuative ARCs. The following diagram sums up the taxonomy:

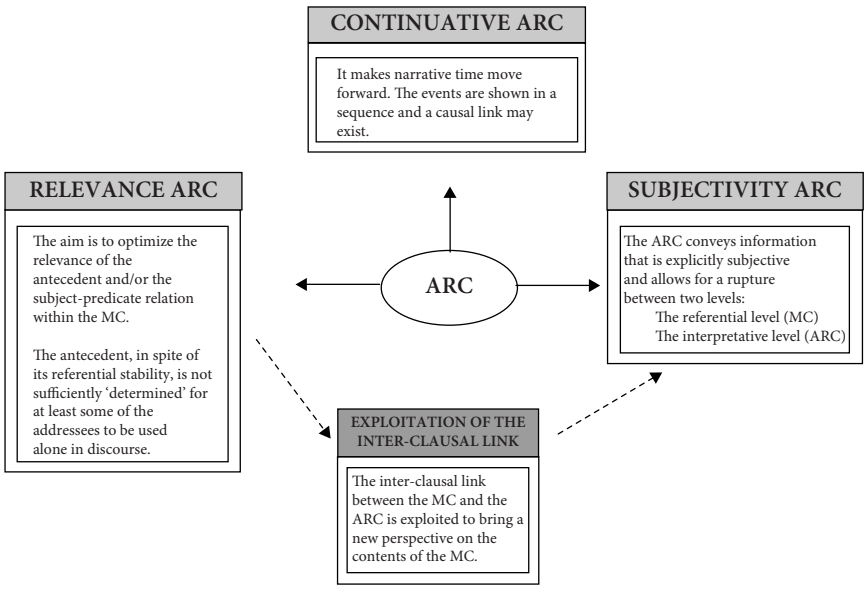


Figure 1. Loock's taxonomy of ARCs

1.1 Relevance ARCs

Relevance ARCs respond to the constraint that a speaker has when s/he needs to convey information known by some of her/his addressees but unknown by some others. Such ARCs fulfil this need for a compromise to ensure that the relevance of the utterance is optimized for no gratuitous mental effort (following Sperber and Wilson's (1986) definition of relevance). (2) illustrates this category:

(2) a. *he was convinced # the battle # for the hearts # and minds of the people # was being won # especially # among the Ovambo # who form the majority # of SWAPO's support*

b. *normally visitors to the state department require credentials # and even then # they have to pass through metal detectors # but twenty year old # Edward Steven Doster # managed to evade the security arrangements # and carry # a collapsible rifle # inside # and up to the seventh floor # where the secretary of state # has his offices*

This type of ARCs encompasses different discursive strategies: (i) levelling of the shared cognitive space (the speaker inserts supplementary information to compensate for the differences in the amount of knowledge shared by the participants), (ii) legitimacy of the antecedent and/or the subject-predicate relation in the MC, or (iii) explanation, justification of or concession in opposition to the information content of the MC, the link being inferred by the addressees (see Loock 2007: 46–50).

1.2 Subjectivity ARCs

Speakers also need to convey with an ARC (rather than an independent clause, for example) information that represents a comment, a judgement or an assessment, by themselves or somebody else. This kind of ARCs, labelled subjectivity ARCs, establishes a discrepancy between a referential level (the main clause) and a commentary level (the ARC). Example (3) below illustrates this category:

(3) a. *Israelis # have sympathy and liking for Americans # which is just as well # since the country is swarming # with transatlantic visitors*

b. *most of them were made of nylon # and imported # which I found very very strange*

1.3 Continuative ARCs

Finally we distinguish continuative ARCs, already defined by Jespersen (1970) and Cornilescu (1981) among others although the definitions are not interchangeable in any systematic way. Such ARCs are used to "make narrative time move forward", i.e. to depict two successive events, quite an unusual discourse function for an embedded clause (Depraetere 1996). A causal link may exist between the MC

and the ARC, the first event triggering the event depicted by the relative clause. Example (4) illustrates this category:

(4) a. *northern Scotland will have occasional light rain # which will be followed during the day by colder but still mainly cloudy weather # with a few sleet and snow showers*
b. *the first book he took from the library was Darwin's # Origin of Species # which inspired him with the dream of becoming a geologist*

This last category is clearly different from the first two categories, in particular regarding the hierarchization of the informational contents. By depicting two successive extra-linguistics events, continuative ARCs are exceptional, as such narrative dynamism is restricted to independent clauses (Depraetere 1996). Therefore, the informational content seems to be on the same level, each being interpreted as belonging to the foreground. This idea paves the way for the suggestion found in the literature that such ARCs share syntactic characteristics with independent clauses and not subordinate clauses. This idea is also expressed for ARCs as a whole, some linguists considering that ARCs are main clauses that are somehow interpolated within a first clause (e.g. Ross 1967; Emonds 1979; McCawley 1982; Fabb 1990). We cannot go in too much detail here about this thorny debate, but we wish to stress the potential interest of investigating the possible prosodic realization of ARCs as independent clauses.

1.4 Morphosyntactic, semantic and prosodic characteristics

These categories can be illustrated with morphosyntactic and semantic phenomena, ARCs in each category displaying particular characteristics. For example, relevance ARCs are most of the time apposed to an antecedent that is the subject in the main clause, while that of a continuative ARCs is generally an object (direct or indirect), or an adjunct, in accordance with the constraint of the organisation of English sentences, in which the subject is canonically in initial position, while a continuative ARC is necessarily in final position. Also, a typical feature of subjectivity ARCs is that they generally are what we call sentential relatives (i.e. relatives the antecedents of which are not NPs but VPs, whole sentences, or even paragraphs) and therefore in final position in the sentence.

Complementing such morphosyntactic and semantic analyses, this paper aims at investigating the prosodic realization of ARCs in relation to Loock's typology. In this perspective, prosodic features of ARCs can be related to the more general issue of parentheticals (cf. for instance Wichmann (2000): 95). In contradiction to traditional descriptions, which restrict parenthesis to a lowering and narrowing in pitch range often coupled with pauses (cf. Armstrong & Ward (1931), Crystal (1969), Cruttenden (1986)), other analyses, often based on corpus data, suggest

that parenthetical items actually display a diversity of prosodic configurations (cf. Bolinger (1989)). As noted by Wichmann (2001) and more recently by Blakemore (2005) for *and*-parentheticals, such diversity might be related to the prosodic marking of discourse or pragmatic distinctions.

2. Prosodic analysis

2.1 Fundamental prosodic conceptions

The term *prosody* is often equated with those linguistic and paralinguistic phenomena related to the melody of speech. Such a position, which can be explained by both historical and technical reasons (cf. Auran (2004): Chapter 5 for an overview on this issue), epitomizes a generalised bias towards tonal issues in prosodic studies. However, numerous researchers (e.g. Couper-Kuhlen & Selting (1996); Ladd (1996); Hirst & Di Cristo (1998)) also advocate a wider conception taking into account not only those aspects related to tone and intonation, but also other elements related to temporal phenomena, intensity and voice quality (cf. for instance Campbell & Mokhtari (2003); Gobl & Ní Chasaide (2003) or Campbell (2004) regarding this last concept).

In this study, we wish to adopt the latter position and build on Di Cristo's (2000) conception of prosody as a macro-system. We thus propose to view prosody as fundamentally grouping four interrelated but independently analysable acoustically rooted systems; these systems respectively concern tonal aspects (tone and intonation, in relation with speech melody), temporal aspects (unit durations and speech rate), intensity (one of the major correlates of loudness) and voice quality (in relation with spectral characteristics of the speech signal), which, due to technical reasons, could not be dealt with in the present study.

2.2 Prosodic representations

Another generalised misconception regarding prosodic analysis concerns the (lack of) distinctions between levels of representation. Following Hirst et al. (2000), we distinguish, for all the prosodic systems mentioned in the preceding section, between four levels: the acoustic level, the phonetic level, and the surface and deep phonological levels. We will focus more particularly on the first two:

– The acoustic level is related to physical characteristics of the speech signal (fundamental frequency or F0, objective durations, global and band-specific intensity, spectral envelope).

– The phonetic level, which factors out the universal (low-level) physical constraints on speech production and perception, retains elements of linguistic significance. For instance, micro-variations in fundamental frequency related to segments (lower F0 for voiced stops, pitch skip on a vowel preceded by a voiceless consonant, *etc.*) are automatically produced by the speaker but absolutely not perceived by the hearer, and should therefore be identified but not taken into account for subsequent linguistic modelling phases (phonological levels).

The phonetic analyses carried out in the present study, and notably those concerning tonal and temporal aspects, rely on this distinction between the acoustic and phonetic levels. In this perspective, a preliminary phase to tonal analysis implied its modelling using the MOMEL algorithm (cf. Hirst & Espesser (1993) ; Hirst et al. (2000)), which aims at factoring out any micro-segmental characteristics (the "micro-prosodic component", cf. Di Cristo & Hirst (1986)). The resulting curve is thus similar to that found on a sequence of entirely sonorant segments and constitutes the "macro-prosodic component", a phonetic construct as opposed to the purely acoustic F0 curve.

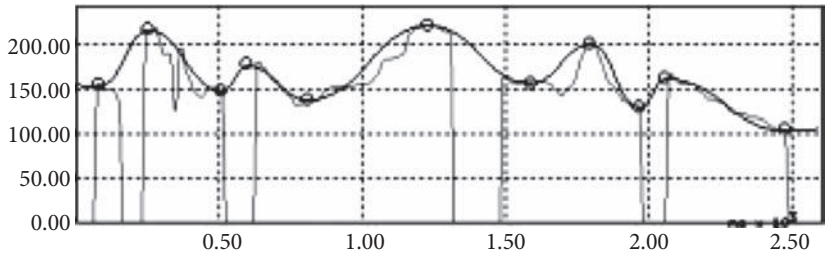


Figure 2. MOMEL modelling of the F0 curve

Temporal representation resorted to z-score transforms, a classical statistical procedure allowing a unit independent (i.e. normalised) representation of data. Normalised durations were thus computed for each phoneme in the corpus (cf. formula 1 below), permitting the neutralisation of specific durational characteristics.

$$norm_dur_{pho_i} = \frac{(actual_dur_{pho_i} - mean_{pho_i})}{sd_{pho_i}}$$

Formula 1. Normalised phoneme duration

For example,

- an occurrence of the phoneme /ə/ (mean duration for this phoneme = 91 ms./ standard deviation = 56 ms.) with an actual duration of 147 ms. would be normalised to norm_dur = 1. With an actual duration of 91 ms. (= mean duration) this phoneme would induce norm_dur = 0.
- and an occurrence of the phoneme /aɪ/ (mean duration for this phoneme = 153 ms./standard deviation = 66 ms.) with an identical actual duration of 147 ms. would be normalised to norm_dur = –0.09

In this particular example, this normalisation method exemplifies the fact that an actual duration of 147 ms. is to be regarded as a lengthening for a /ə/ (positive value), but as a slight shortening for /aɪ/ (negative value), which is inherently longer than /ə/.

2.3 Prosodic dimensions

We shall conclude this brief parsing of fundamental prosodic concepts by mentioning a distinction between two types of dimensions within prosodic systems. Indeed, if the linear succession of F0 ups and downs does constitute a key element of the tonal aspects of prosody, another dimension, which Ladd (1996) calls "orthogonal" is also to be taken into account. Indeed, as far as the tonal system is concerned, the traditional linear (or "horizontal") succession of high and low tones (for instance in the ToBI system; cf. Beckman & Ayers (1994)) actually takes place within a broader ("vertical") frame related to the speaker's register. This concept of register can be divided into two parameters: register (or "pitch") level and span. Figures 3a and 3b below illustrate these concepts.

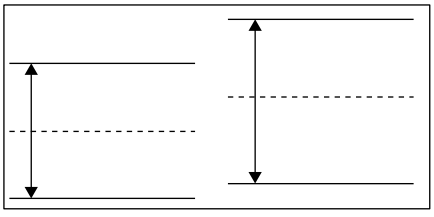


Figure 3a. Difference in register levels

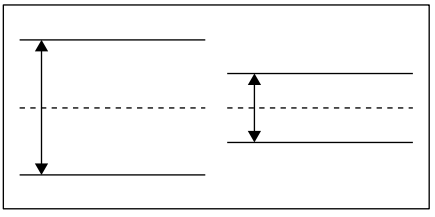


Figure 3b. Difference in register span

We already noted in Section 1.4 that such modifications in register level and span are traditionally associated with parentheticals (prosodic compression) and are thus to be taken into account in the prosodic characterisation of ARCs. Intensity, which can be represented phonetically using methods similar to

those used for tonal aspects (intensity curves), is described as following similar patterns and therefore received particular attention within this study (in spite of difficulties inherent to speaker-to-microphone distance variations, which, though non-linguistic in nature, do influence measurement accuracy).

The temporal system displays a similar distinction whereby the duration of individual linguistic units takes place within the framework of a given speech rate. Among the scarce systematic studies of speech rate variations in relation with discourse (more particularly informational and topical) structure, Koopmans-van Beinum & van Donzel (1996) and Smith (2004), mention a slowing down of speech rate at the beginning of new topical units (paratones). Consequently, the importance of information status in Loock's taxonomy naturally leads us to expect close interactions with speech rate.

2.4 Data extraction

2.4.1 *Discourse annotation*

As mentioned above (Section 0.2), prosodic annotation took place after the identification and analysis of a number of ARCs (discourse annotation). Discourse annotation resulted in the identification of 50 ARCs:

Table 1. Number of items per ARC type

ARC type	Number of items
Relevance	33
Subjectivity	8
Continuative	1
Relevance/Subjectivity	4
Ambiguous continuative	2
Unidentified	2

The unavailability of the ICE-GB, more spontaneous than the Aix-MARSEC, leads to an obvious over-representation of relevance ARCs. In this study, cross-type comparisons only involved relevance and subjectivity ARCs, other types not being sufficiently represented in our data for any sound analysis to be carried out. It is moreover important to note that, due to the very limited number of subjectivity ARCs, formal statistical comparisons between those two ARCs types could not be carried out either.

Each item was annotated using a set of five discourse parameters: ARC type, position (initial, medial or final), information status of the antecedent, information status of the ARC and phrastic status of the antecedent.

2.4.2 *Prosodic annotation*

All ARCs were subsequently semi-automatically analysed using specific scripts within Praat (cf. Boersma (2001); Boersma & Weenink (2006)). The procedure comprises the following steps: automatic loading of the sound file; manual selection of the ARC; manual selection of the ARC onset and offset; manual selection of the previous intonation unit (IU) and its offset; manual selection of the next IU and its onset.

F0 values were normalized using a logarithmic scale (in semitones) in order to allow relevant comparison between speakers.

For each ARC, a total of 48 prosodic parameters were then automatically computed:

- Tonal system (32): ARC mean F0 (Htz + semitones or ST), ARC minimum F0 (Htz + ST), ARC maximum F0 (Htz + ST), ARC register span (Htz + ST), ARC onset (Htz + ST), ARC offset (Htz + ST), previous IU mean F0 (Htz + ST), previous IU minimum F0 (Htz + ST), previous IU maximum F0 (Htz + ST), previous IU register span (Htz + ST), previous IU offset (Htz + ST), next IU mean F0 (Htz + ST), next IU minimum F0 (Htz + ST), next IU maximum F0 (Htz + ST), next IU register span (Htz + ST), next IU onset (Htz + ST), difference between previous IU offset and ARC onset (ST), difference between ARC offset and next IU onset (ST)
- Temporal system (10): ARC duration (raw and normalised), previous IU duration (raw and normalised), next IU duration (raw and normalised), difference between previous IU normalised duration and ARC normalised duration, difference between ARC normalised duration and next IU normalised duration, silence duration before ARC, silence duration after ARC
- Intensity system (6): mean of ARC global intensity, standard deviation of ARC global intensity, mean of previous IU global intensity, standard deviation of previous IU global intensity, mean of next IU global intensity, standard deviation of next IU global intensity

A total of 2173 observations (41 ARCs × (5 discourse parameters + 48 prosodic parameters)) were then fed into the R software for statistical analysis.

3. Results

3.1 ARCs as parentheticals

As far as tonal aspects are concerned, we investigated register level and span, onset value and its difference with the offset of the preceding IU and offset value and the

difference with the onset of the following IU. The register level of ARCs was found to be significantly lower than that of both the preceding (Kolmogorov-Smirnov test; $p = 0.001 < 0.05$) and the following (Kolmogorov-Smirnov test; $p = 0.012 < 0.05$) IU; this behaviour is recognised as typical of prosodic parentheticals. Other indicators, however, were found to diverge from typical parentheticals: first, ARC register span was not identified as significantly different from both preceding and following IUs (Kolmogorov-Smirnov test; $p = 0.964 > 0.05$ and $p = 0.711 > 0.05$); second, ARC onset differential, though not significantly different from that of the following IU (Kolmogorov-Smirnov test; $p = 0.396 > 0.05$), displayed an unusual positive value (mean = 2.24 ST), commonly associated with discourse discontinuity (cf. Auran 2004). Differences across ARC types will be dealt with in Section 3.2.

Temporal analyses showed that no significant differences could be found between speech rates for ARCs and their surrounding IUs (Kolmogorov-Smirnov test; $p = 0.259 > 0.05$ and $p = 0.380 > 0.05$). However, these results seem to neutralise differences related to parameters such as antecedent type (Kolmogorov-Smirnov test; $p = 0.0004 < 0.05$), which will be explored in future work.

The analysis of intensity parameters (level and span) revealed no significant differences between ARCs and their surrounding IUs (mean: Kolmogorov-Smirnov test; $p = 0.068 > 0.05$ and $p = 0.179 > 0.05$/standard deviation: Kolmogorov-Smirnov test; $p = 0.396 > 0.05$ and $p = 0.179 > 0.05$).

These results suggest a complex interplay of production and interpretation constraints whereby ARCs show characteristics both traditional (register level, intensity level) and atypical (register span, speech rate and intensity span) of parentheticals.

3.2 Differences between types of ARCs

For reasons related to a limited number of subjectivity ARCs, the results presented here reflect but tendencies, which would have to be confirmed through formal statistical testing. These preliminary results, however, seem to indicate prosodic differences that can be interpreted as differences in discourse functions. Indeed, in spite of a lack of clear tendencies concerning both register level and span, a stark dichotomy can be drawn between those ARC types, with apparently higher onset values for subjectivity ARCs (mean normalised values: relevance ARCs = 1.80 ST above speaker's average/subjectivity ARCs = 2.23 ST above speaker's average). This can clearly be interpreted as a sign of stronger discourse discontinuity for subjectivity ARCs (cf. Brown & Yule 1983; Wichmann 2000; Auran 2004).

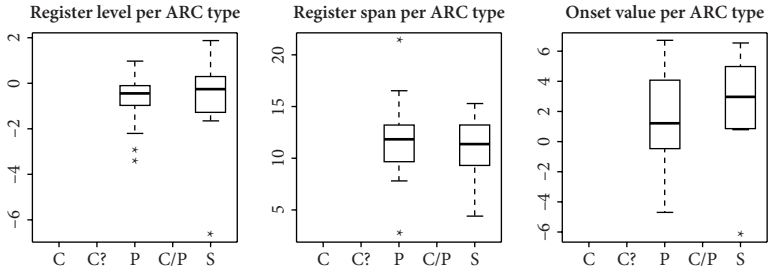


Figure 4. Comparative results between relevance and subjectivity ARCs (intensity and temporal aspects)

Intensity results show no differences regarding span but do signal lower level values for subjectivity ARCs, which seems surprising given the involvement traditionally associated with subjectivity (Caelen & Auran (2004)).

Speech rate measurements seem to indicate a clear-cut difference between relevance and subjectivity ARCs, the latter being characterised by longer normalised durations corresponding to a slower rate (relevance ARC mean normalised duration = –0.178/subjectivity ARC mean normalised duration = –0.043).

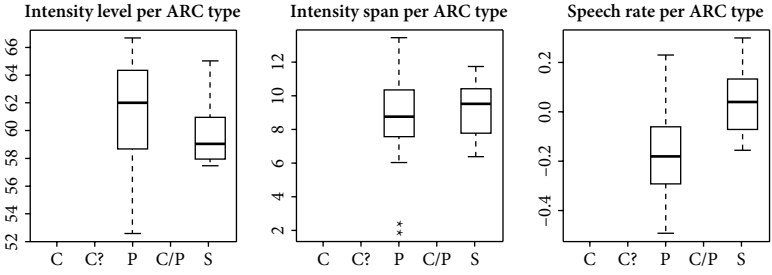


Figure 5. Comparative results between relevance and subjectivity ARCs (intensity and temporal aspects)

Table 2 summarises the tonal, temporal and intensity results for all ARCs and for each ARC type.

Table 2. Summary of main results (tendencies in relative terms) [low = –/high = +,++/neutral = 0]

	Tonal parameters			**Temporal parameters**	**Intenstity parameters**	
	Level	**Span**	**Δ Onset**	**Speech rate**	**Level**	**Span**
All ARCs	–	0	+	0	0	0
Relevance ARCs	–	0	+	0	0	0
Subjectivity ARCs	–	0	++	Slow	–	0

4. Discussion

The surprisingly atypical characteristics of ARCs as a whole seem to go along with the idea that ARCs may have the syntactic behaviour and the semantic interpretation of independent clauses (cf. Section 1.3). Register and intensity levels (particularly for subjectivity ARCs), both lower than those of surrounding units, are characteristic of prosodic parentheticals; but their register and intensity spans, together with their speech rate, clearly correspond to classical IUs realizing independent clauses. It may be interesting to relate this fact with the possibility for ARCs to convey independent speech acts (cf. Emonds 1979; McCawley 1982 among others).

The most striking phenomenon regarding the distinction between relevance and subjectivity ARCs concerns discourse discontinuity marking through high onset values for both types; subjectivity ARCs display even stronger discontinuity, which seems in line with a more important rupture with the discourse topic corresponding to a shift between the referential and interpretative levels mentioned in figure 1. Although some rupture is present for both types, the information conveyed by a subjectivity ARC is somehow more peripheral: it does not provide any information to optimise the relevance of the antecedent or/and the contents of the main clause, or information to fill in a supposedly gap in (some of) the addressees' knowledge, but a non-topical comment or judgement.

The lower intensity level values measured for subjectivity ARCs may, at first sight, seem somehow counter-intuitive, but can easily be explained if we consider the fact that subjective episodes in discourse often display apparently conflicting prosodic characteristics (cf. Di Cristo et al. (2004)); this can constitute a strategy used by the speaker to induce the perception of intermediate levels between otherwise discrete categories such as continuity/discontinuity, subjectivity/objectivity, etc. More specifically, reduced intensity parameters (compression) are often used in synchrony with increased tonal parameters (expansion), thus conveying an intermediate level of personal involvement in the discourse at stake.

The observed clear-cut difference in speech rates, eventually, may not be analysed only in terms of discourse functions, since the great majority of subjectivity ARCs qualifies sentential antecedents (cf. Loock (2007)); those two parameters (subjectivity and syntactic nature of the antecedent) are therefore difficult to separate. The investigation of the respective influence of both these parameters requires further research, which constitutes a forthcoming phase of our project.

5. Conclusion

Within our global project dealing with form-function relations in spoken discourse, this preliminary study clearly shows that various discourse functions associated with a given syntactic structure give way to differences in prosodic realization. Not only have we provided evidence in favour of a view of ARCs as atypical parentheticals, but we have also proposed that prosodic markers can serve as input constraints influencing the pragmatic interpretation of one syntactic structure in discourse.

Although it was restricted to the analysis of two of Loock's ARC categories, this work also questions the traditional boundary between independent and embedded clauses, for which continuative ARCs particularly are described as problematic. Further research, extending the methodology used here, will tackle this issue and allow a closer description of the prosodic characteristics of ARCs in relation to their discourse functions.

References

Armstrong, L. & Ward, I. 1931. *A Handbook of English Intonation*, Cambridge, Heffer.

Auran, C. 2004. *Prosodie et anaphore dans le discours en anglais et en français: cohesion et attribution référentielle*. Ph.D. Dissertation, Université de Provence, France and Laboratoire Parole et Langage UMR 6057, CNRS.

Auran, C., Bouzon, C. & Hirst, D.J. 2004. "The Aix-MARSEC project: an evolutive database of spoken British English", *Speech Prosody 2004*, Nara, 561–564.

Beckman, M.E. & Ayers, G.M. 1994. *ToBI annotation conventions*. http://ling.ohio-state.edu/-tobi/ame_tobi.

Blakemore, D. 2005. "*And*-parentheticals", in *Journal of Pragmatics* 37, 1165–1181.

Boersma, P. 2001. "Praat, a system for doing phonetics by computer", *Glot International* 5:9/10, 341–345.

Boersma, P. & Weennink, D. 2006. *Praat: doing phonetics by computer (Version 4.4.17)* [Computer program]. Retrieved April 19, 2006, from http://www.praat.org/.

Bolinger, D. 1989. *Intonation and Its Uses*, London, Edward Arnold.

Brown, G. & Yule, G. 1983. *Discourse analysis*. Cambridge, C.U.P.

Caelen, G. & Auran, C. 2004. "The Phonology of Melodic Prominence: the structure of melisms", in *Speech Prosody 2004*, Nara, 143–146.

Campbell, N. 2004. "Accounting for voice-quality variation", in *Speech Prosody 2004*, Nara, 217–220.

Campbell, N. & Mokhtari 2003. "Voice quality: the 4th prosodic dimension", in 15th ICPhS (ICPhS'03), Barcelona, Spain, 2417–2420.

Cornilescu, A. 1981. "Non-restrictive Relative Clauses, an Essay in Semantic Description", in *Revue roumaine de linguistique* XXVI, 1, 41–67.

Couper-Kuhlen, E. & Selting, M. 1996. *Prosody in Conversation. Interactional Studies*, Cambridge, C.U.P.

Cruttenden, A. 1986. *Intonation*, Cambridge, C.U.P.
Depraetere, I. 1996. "Foregrounding in English Relative Clauses", *Linguistics* 34, 699–731.
Di Cristo, A. 2000. "La problématique de la prosodie dans l'étude de la parole dite spontanée", in *Revue Parole 15–16*, 189–250.
Di Cristo, A., Auran C., Bertrand R., Chanet C., Portes C., Régnier A. 2004. "Outils prosodiques et analyse du discours", in *CILL 30 (1–3)*, 27–84.
Di Cristo, A. & Hirst, D.J. 1986. "Modelling French micromelody: analysis and synthesis", in *Phonetica* 43, 1–3, 11–30.
Di Cristo, A. & Hirst, D.J. 1998. *Intonation Systems: A survey of Twenty Languages*, Cambridge, C.U.P.
Emonds, J. 1979. "Appositive Relatives Have No Properties", in *Linguistic Inquiry* 10.2, 241–3.
Fabb, N. 1990. "The Difference between English Restrictive and Nonrestrictive Relative Clauses", in *Linguistics* 26, 57–78.
Gobl, C. & Ní Chasaide, A. 2003. "The role of voice quality in communicating emotion, mood and attitude", in *Speech Communication* 40, 189–212.
Greenbaum, S. (ed.). 1996. *Comparing English Worldwide: The International Corpus of English*. Oxford, Clarendon Press.
Hirst, D.J. & Espesser, R. 1993. "Automatic modelling of fundamental frequency using a quadratic spline function", in *Travaux de l'Institut de Phonétique d'Aix* 15, 71–85
Hirst, D.J., Di Cristo, A. & Espesser, R. 2000. "Levels of Representation and Levels of Analysis for the Description of Intonation Systems", *in* Horne, M. (ed.), *Prosody: Theory and Experiment. Text, Speech and Language Technbology*, 14. Kluwer Academic Publishers, 51–87.
Jespersen, O. [1927] 1970. *A Modern English Grammar on Historical Principles.* Vol III. London, George Allen & Unwin.
Knowles, G., Wichmann, A. & Alderson, P. 1996. *Working with Speech: perspectives on research into the Lancaster/IBM Spoken English Corpus*. London, Longman.
Koopmans-Van Beinum, F.J.& Donzel Van, M.E. 1996. "Relationship between discourse structure and dynamic speech rate", in *Proceedings ICSLP96, Fourth ICSLP*, *Vol 3*, Philadelphia, 1724–1727.
Ladd, R. 1996. *Intonational Phonology*. Ccambridge, C.U.P.
Loock, R. 2003. "Les Fonctions discursives des propositions subordonnées relatives 'appositives' en discours", in *Anglophonia* 12, 113–31.
Loock, R. 2005. *La Proposition subordonnée relative appositive en anglais contemporain à l'écrit et à l'oral: fonctions discursives et structures concurrentes*. Ph.D. Dissertation, Lille III University, France.
Loock, R. 2007. "Appositive Relative Clauses and their Functions in Discourse", in *Journal of Pragmatics 39*: 336–62.
McCawley, J.D. 1982. "Parentheticals and discontinuous constituent structure", in *Linguistic Inquiry* 13.1, 91–106.
Prince, E F. 1981. "Toward a Taxonomy of Given/New Information", in *Radical Pragmatics,* Peter COLE, ed. New York: Academic Press, pp 223–54.
Prince, E.F. 1986. "On the Syntactic Marking of Presupposed Open Propositions", in Farley, A. et al. (eds), *Papers from the Parasession on Pragmatics and Grammatical Theory,* 22nd regional meeting of the Chicago Linguistics Society, 208–22.
Prince, E.F. 1992. "The ZPG Letter: Subjects, Definiteness, and Information-Status", in Mann, William C and Thompson, Sandra A. (eds), *Discourse Description: Diverse Analyses of a Fund Raising Text*. Philadelphia, John Benjamins B.U., 295–325.

R Language and Environment for Statistical Modelling: available from http://lib.stat.cmu.edu/R/CRAN/.

Roach, P., Knowles, G., Varadi, T. & Aenfield, S. 1993. "MARSEC: A machine readable Spoken English corpus", in *Journal of the International Phonetic Association* 23: 2, 47–53.

Ross, J. 1967, *Constraints on Variables in Syntax*, Ph.D. Dissertation, MIT.

Smith, C. 2004. "Topic transitions and durational prosody in reading aloud: production and modeling", in *Speech Communication 42*, 247–270.

Sperber, D. & Wilson, D. 1986. *Relevance, communication and cognition*. Oxford, Blackwell Publishers.

Wichmann, A. 2000. *Intonation in Text and Discourse*. London, Longman.

Wichmann, A. 2001. "Spoken parentheticals", *in* Aijmer, K. (ed.), *A Wealth of English: Studies in Honour of Goran Kjellmer*. Gothenburg, Gothenburg University Press, 171–193.

Index

A
adjacency 15, 17–23
anaphora 8, 27, 43–44, 141–142, 144–157, 159–160
annotation 81, 86–87, 89, 91, 94, 96, 102–103, 164, 171–172

B
bidirectionality 36–39, 55

C
chain graph 9
clause 15–17, 21–25, 27, 29–30, 32–33, 143–144, 166–167, 175
coherence 1, 5, 15–16, 21–25, 33, 47, 63
cohesion 1
completion 71–74, 89
commissive 130
constraint 35–41, 45–46, 67, 89, 92, 107–109, 142–144, 155, 159
  soft constraint 46
  statistic constraint 93
construction 4, 113, 115
context 40–41, 55, 65, 77, 109, 117–121, 127–129, 132, 135, 137, 147–148, 154–155
CONTEXT 132, 135, 137
  minimal context 127–128, 137
contextual parameter 55
convention 128, 131–137
convention of usage 128, 134, 137
conventionalized metonymy 134
coreference 18, 21–23, 25–34, 62–64, 68–71, 73–74, 76–77, 141–142
corpus 81–82, 84–85, 95, 99, 126, 164
correct 3, 11
cue 84, 87, 92
Cypriot Greek 126, 129

D
directed acyclic graph (DAG) 9
discourse coherence relation 82
discourse constituency unit (DCU) 43–45, 49–54
discourse markers 45, 62, 83–84, 103
discourse parser 81, 102–103
discourse purpose 3
discourse referent 56, 145–151, 153, 157–158
discourse segment 3, 81–82, 88, 102
  complex discourse segment (CDS) 82, 84, 88–89, 93, 95
  elementary discourse segment (EDS) 84, 95
dual response 135–136
dual uptake 135–136

E
event 19, 28, 31, 52–53, 62–63, 68–71, 73–74, 76–77, 151
eventuality 4–5, 16, 20, 26–28, 116
explanation 52, 63, 75–77
  best explanation 17–21, 33

F
F0 3, 74, 168–170, 172

H
head driven phrase structure grammar (HPSG) 131–132, 135, 137, 159

I
idiom 126, 132
illocutionary force 125–126, 128, 136
  conventional illocutionary force 125–126, 132–135, 137
illocutionary force indicating device (IFID) 131
implicature 48, 107–110, 117–119, 121
inference 62–64, 69, 71, 132–133
intention 3–6, 38, 40–41, 47, 82–83, 133
interpretation 17–22, 29–30, 33–34, 36–46, 62–65, 67, 69–71, 126, 129–130, 146, 155–156, 160, 176
  formulaic interpretation 129, 135
interpretive shortcut 126, 137
intonation 2, 39, 46, 62, 70–75, 113, 116, 132, 168, 172–173

L
lexicalization 129, 131

M
minimal unit 2–5
nuclearity 6–7
mono-nuclear 6–7
multi-nuclear 6–7

O
offer 127, 131, 133–135
  conventional offer 134–135
optimality theory (OT) 35–38, 40–41, 56–57

P
performative hypothesis 137
plan-based approach 128
precondition 134–135
presupposition 105–106, 110–122, 150, 157–158
proof 17–20, 22, 29–31, 33

proposition 42, 50, 52, 114, 116
prosody 2–3, 163, 167–168, 170, 172, 175–176

R
relative clause 17, 163
  appositive relative clause (ARC) 163–167, 171–176
  determinative relative clause (DRC) 165
request 127, 129, 131, 133–136
rhetorical relation 5–8, 49–50, 53–54, 85, 95–96
rhetorical structure 1–10, 43–44, 54–56, 82, 92
Rhetorical Structure Theory (RST) 4, 7–10, 20, 43, 83, 85, 92–93, 96–101
Right Frontier Constraint(RFC) 7–8, 54

S
scientific article 84–85, 91, 93, 95, 99
Segmented Discourse Representation Theory (SDRT) 5–10, 63, 120
semantics, compositional 126, 132–136
semantics, interrogative 136
speech 40, 47, 49, 62, 2–74, 127, 128–129, 158, 163, 168–169, 175
speech act 49, 74, 125–126, 137
speech act formulae 125–126, 129–131, 137
speech rate 168, 171, 173–175
state 19, 27, 31

T
text 1, 17–18, 43, 88–93
text generation 43
text type structure (TTS) 81–82, 84, 90–102
  logical document structure 81–82, 84–89, 102
  generic document structure 81–82, 84–85, 90–102
transcription 164
tree 5, 9–10, 54, 89

U
uptake 127

Pragmatics & Beyond New Series

A complete list of titles in this series can be found on *www.benjamins.com*

197 **DEDAIC, Mirjana N. and Mirjana MISKOVIC-LUKOVIC (eds.):** South Slavic Discourse Particles. ix, 162 pp. + index. *Expected June 2010*

196 **STREECK, Jürgen (ed.):** New Adventures in Language and Interaction. vi, 269 pp. + index. *Expected June 2010*

195 **PAHTA, Päivi, Minna NEVALA, Arja NURMI and Minna PALANDER-COLLIN (eds.):** Social Roles and Language Practices in Late Modern English. viii, 227 pp. + index. *Expected June 2010*

194 **KÜHNLEIN, Peter, Anton BENZ and Candace L. SIDNER (eds.):** Constraints in Discourse 2. 2010. v, 180 pp.

193 **SUOMELA-SALMI, Eija and Fred DERVIN (eds.):** Cross-Linguistic and Cross-Cultural Perspectives on Academic Discourse. 2009. vi, 299 pp.

192 **FILIPI, Anna:** Toddler and Parent Interaction. The organisation of gaze, pointing and vocalisation. 2009. xiii, 268 pp.

191 **OGIERMANN, Eva:** On Apologising in Negative and Positive Politeness Cultures. 2009. x, 296 pp.

190 **FINCH, Jason, Martin GILL, Anthony JOHNSON, Iris LINDAHL-RAITTILA, Inna LINDGREN, Tuija VIRTANEN and Brita WÅRVIK (eds.):** Humane Readings. Essays on literary mediation and communication in honour of Roger D. Sell. 2009. xi, 160 pp.

189 **PEIKOLA, Matti, Janne SKAFFARI and Sanna-Kaisa TANSKANEN (eds.):** Instructional Writing in English. Studies in honour of Risto Hiltunen. 2009. xiii, 240 pp.

188 **GILTROW, Janet and Dieter STEIN (eds.):** Genres in the Internet. Issues in the theory of genre. 2009. ix, 294 pp.

187 **JUCKER, Andreas H. (ed.):** Early Modern English News Discourse. Newspapers, pamphlets and scientific news discourse. 2009. vii, 227 pp.

186 **CALLIES, Marcus:** Information Highlighting in Advanced Learner English. The syntax–pragmatics interface in second language acquisition. 2009. xviii, 293 pp.

185 **MAZZON, Gabriella:** Interactive Dialogue Sequences in Middle English Drama. 2009. ix, 228 pp.

184 **STENSTRÖM, Anna-Brita and Annette Myre JØRGENSEN (eds.):** Youngspeak in a Multilingual Perspective. 2009. vi, 206 pp.

183 **NURMI, Arja, Minna NEVALA and Minna PALANDER-COLLIN (eds.):** The Language of Daily Life in England (1400–1800). 2009. vii, 312 pp.

182 **NORRICK, Neal R. and Delia CHIARO (eds.):** Humor in Interaction. 2009. xvii, 238 pp.

181 **MASCHLER, Yael:** Metalanguage in Interaction. Hebrew discourse markers. 2009. xvi, 258 pp.

180 **JONES, Kimberly and Tsuyoshi ONO (eds.):** Style Shifting in Japanese. 2008. vii, 335 pp.

179 **SIMÕES LUCAS FREITAS, Elsa:** Taboo in Advertising. 2008. xix, 214 pp.

178 **SCHNEIDER, Klaus P. and Anne BARRON (eds.):** Variational Pragmatics. A focus on regional varieties in pluricentric languages. 2008. vii, 371 pp.

177 **RUE, Yong-Ju and Grace ZHANG:** Request Strategies. A comparative study in Mandarin Chinese and Korean. 2008. xv, 320 pp.

176 **JUCKER, Andreas H. and Irma TAAVITSAINEN (eds.):** Speech Acts in the History of English. 2008. viii, 318 pp.

175 **GÓMEZ GONZÁLEZ, María de los Ángeles, J. Lachlan MACKENZIE and Elsa M. GONZÁLEZ ÁLVAREZ (eds.):** Languages and Cultures in Contrast and Comparison. 2008. xxii, 364 pp.

174 **HEYD, Theresa:** Email Hoaxes. Form, function, genre ecology. 2008. vii, 239 pp.

173 **ZANOTTO, Mara Sophia, Lynne CAMERON and Marilda C. CAVALCANTI (eds.):** Confronting Metaphor in Use. An applied linguistic approach. 2008. vii, 315 pp.

172 **BENZ, Anton and Peter KÜHNLEIN (eds.):** Constraints in Discourse. 2008. vii, 292 pp.

171 **FÉLIX-BRASDEFER, J. César:** Politeness in Mexico and the United States. A contrastive study of the realization and perception of refusals. 2008. xiv, 195 pp.

170 **OAKLEY, Todd and Anders HOUGAARD (eds.):** Mental Spaces in Discourse and Interaction. 2008. vi, 262 pp.

169 **CONNOR, Ulla, Ed NAGELHOUT and William ROZYCKI (eds.):** Contrastive Rhetoric. Reaching to intercultural rhetoric. 2008. viii, 324 pp.

168 **PROOST, Kristel:** Conceptual Structure in Lexical Items. The lexicalisation of communication concepts in English, German and Dutch. 2007. xii, 304 pp.

167 **BOUSFIELD, Derek:** Impoliteness in Interaction. 2008. xiii, 281 pp.

166 **NAKANE, Ikuko:** Silence in Intercultural Communication. Perceptions and performance. 2007. xii, 240 pp.

165 **BUBLITZ, Wolfram and Axel HÜBLER (eds.):** Metapragmatics in Use. 2007. viii, 301 pp.

164 **ENGLEBRETSON, Robert (ed.):** Stancetaking in Discourse. Subjectivity, evaluation, interaction. 2007. viii, 323 pp.

163 **LYTRA, Vally:** Play Frames and Social Identities. Contact encounters in a Greek primary school. 2007. xii, 300 pp.

162 **FETZER, Anita (ed.):** Context and Appropriateness. Micro meets macro. 2007. vi, 265 pp.

161 **CELLE, Agnès and Ruth HUART (eds.):** Connectives as Discourse Landmarks. 2007. viii, 212 pp.

160 **FETZER, Anita and Gerda Eva LAUERBACH (eds.):** Political Discourse in the Media. Cross-cultural perspectives. 2007. viii, 379 pp.

159 **MAYNARD, Senko K.:** Linguistic Creativity in Japanese Discourse. Exploring the multiplicity of self, perspective, and voice. 2007. xvi, 356 pp.

158 **WALKER, Terry:** *Thou* and *You* in Early Modern English Dialogues. Trials, Depositions, and Drama Comedy. 2007. xx, 339 pp.

157 **CRAWFORD CAMICIOTTOLI, Belinda:** The Language of Business Studies Lectures. A corpus-assisted analysis. 2007. xvi, 236 pp.

156 **VEGA MORENO, Rosa E.:** Creativity and Convention. The pragmatics of everyday figurative speech. 2007. xii, 249 pp.

155 **HEDBERG, Nancy and Ron ZACHARSKI (eds.):** The Grammar–Pragmatics Interface. Essays in honor of Jeanette K. Gundel. 2007. viii, 345 pp.

154 **HÜBLER, Axel:** The Nonverbal Shift in Early Modern English Conversation. 2007. x, 281 pp.

153 **ARNOVICK, Leslie K.:** Written Reliquaries. The resonance of orality in medieval English texts. 2006. xii, 292 pp.

152 **WARREN, Martin:** Features of Naturalness in Conversation. 2006. x, 272 pp.

151 **SUZUKI, Satoko (ed.):** Emotive Communication in Japanese. 2006. x, 234 pp.

150 **BUSSE, Beatrix:** Vocative Constructions in the Language of Shakespeare. 2006. xviii, 525 pp.

149 **LOCHER, Miriam A.:** Advice Online. Advice-giving in an American Internet health column. 2006. xvi, 277 pp.

148 **FLØTTUM, Kjersti, Trine DAHL and Torodd KINN:** Academic Voices. Across languages and disciplines. 2006. x, 309 pp.

147 **HINRICHS, Lars:** Codeswitching on the Web. English and Jamaican Creole in e-mail communication. 2006. x, 302 pp.

146 **TANSKANEN, Sanna-Kaisa:** Collaborating towards Coherence. Lexical cohesion in English discourse. 2006. ix, 192 pp.

145 **KURHILA, Salla:** Second Language Interaction. 2006. vii, 257 pp.

144 **BÜHRIG, Kristin and Jan D. ten THIJE (eds.):** Beyond Misunderstanding. Linguistic analyses of intercultural communication. 2006. vi, 339 pp.

143 **BAKER, Carolyn, Michael EMMISON and Alan FIRTH (eds.):** Calling for Help. Language and social interaction in telephone helplines. 2005. xviii, 352 pp.

142 **SIDNELL, Jack:** Talk and Practical Epistemology. The social life of knowledge in a Caribbean community. 2005. xvi, 255 pp.

141 **ZHU, Yunxia:** Written Communication across Cultures. A sociocognitive perspective on business genres. 2005. xviii, 216 pp.

140 **BUTLER, Christopher S., María de los Ángeles GÓMEZ GONZÁLEZ and Susana M. DOVAL-SUÁREZ (eds.):** The Dynamics of Language Use. Functional and contrastive perspectives. 2005. xvi, 413 pp.

139 **LAKOFF, Robin T. and Sachiko IDE (eds.):** Broadening the Horizon of Linguistic Politeness. 2005. xii, 342 pp.

138 **MÜLLER, Simone:** Discourse Markers in Native and Non-native English Discourse. 2005. xviii, 290 pp.

137 **MORITA, Emi:** Negotiation of Contingent Talk. The Japanese interactional particles *ne* and *sa*. 2005. xvi, 240 pp.

136 **SASSEN, Claudia:** Linguistic Dimensions of Crisis Talk. Formalising structures in a controlled language. 2005. ix, 230 pp.

135 **ARCHER, Dawn:** Questions and Answers in the English Courtroom (1640–1760). A sociopragmatic analysis. 2005. xiv, 374 pp.

134 **SKAFFARI, Janne, Matti PEIKOLA, Ruth CARROLL, Risto HILTUNEN and Brita WÅRVIK (eds.):** Opening Windows on Texts and Discourses of the Past. 2005. x, 418 pp.

133 **MARNETTE, Sophie:** Speech and Thought Presentation in French. Concepts and strategies. 2005. xiv, 379 pp.

132 **ONODERA, Noriko O.:** Japanese Discourse Markers. Synchronic and diachronic discourse analysis. 2004. xiv, 253 pp.

131 **JANOSCHKA, Anja:** Web Advertising. New forms of communication on the Internet. 2004. xiv, 230 pp.

130 **HALMARI, Helena and Tuija VIRTANEN (eds.):** Persuasion Across Genres. A linguistic approach. 2005. x, 257 pp.

129 **TABOADA, María Teresa:** Building Coherence and Cohesion. Task-oriented dialogue in English and Spanish. 2004. xvii, 264 pp.

128 **CORDELLA, Marisa:** The Dynamic Consultation. A discourse analytical study of doctor–patient communication. 2004. xvi, 254 pp.

127 **BRISARD, Frank, Michael MEEUWIS and Bart VANDENABEELE (eds.):** Seduction, Community, Speech. A Festschrift for Herman Parret. 2004. vi, 202 pp.

126 **WU, Yi'an:** Spatial Demonstratives in English and Chinese. Text and Cognition. 2004. xviii, 236 pp.

125 **LERNER, Gene H. (ed.):** Conversation Analysis. Studies from the first generation. 2004. x, 302 pp.

124 **VINE, Bernadette:** Getting Things Done at Work. The discourse of power in workplace interaction. 2004. x, 278 pp.

123 **MÁRQUEZ REITER, Rosina and María Elena PLACENCIA (eds.):** Current Trends in the Pragmatics of Spanish. 2004. xvi, 383 pp.

122 **GONZÁLEZ, Montserrat:** Pragmatic Markers in Oral Narrative. The case of English and Catalan. 2004. xvi, 410 pp.

121 **FETZER, Anita:** Recontextualizing Context. Grammaticality meets appropriateness. 2004. x, 272 pp.

120 **AIJMER, Karin and Anna-Brita STENSTRÖM (eds.):** Discourse Patterns in Spoken and Written Corpora. 2004. viii, 279 pp.

119 **HILTUNEN, Risto and Janne SKAFFARI (eds.):** Discourse Perspectives on English. Medieval to modern. 2003. viii, 243 pp.

118 **CHENG, Winnie:** Intercultural Conversation. 2003. xii, 279 pp.

117 **WU, Ruey-Jiuan Regina:** Stance in Talk. A conversation analysis of Mandarin final particles. 2004. xvi, 260 pp.

116 **GRANT, Colin B. (ed.):** Rethinking Communicative Interaction. New interdisciplinary horizons. 2003. viii, 330 pp.

115 **KÄRKKÄINEN, Elise:** Epistemic Stance in English Conversation. A description of its interactional functions, with a focus on *I think*. 2003. xii, 213 pp.

114 **KÜHNLEIN, Peter, Hannes RIESER and Henk ZEEVAT (eds.):** Perspectives on Dialogue in the New Millennium. 2003. xii, 400 pp.

113 **PANTHER, Klaus-Uwe and Linda L. THORNBURG (eds.):** Metonymy and Pragmatic Inferencing. 2003. xii, 285 pp.

112 **LENZ, Friedrich (ed.):** Deictic Conceptualisation of Space, Time and Person. 2003. xiv, 279 pp.

111 **ENSINK, Titus and Christoph SAUER (eds.):** Framing and Perspectivising in Discourse. 2003. viii, 227 pp.

110 **ANDROUTSOPOULOS, Jannis K. and Alexandra GEORGAKOPOULOU (eds.):** Discourse Constructions of Youth Identities. 2003. viii, 343 pp.

109 **MAYES, Patricia:** Language, Social Structure, and Culture. A genre analysis of cooking classes in Japan and America. 2003. xiv, 228 pp.

108 **BARRON, Anne:** Acquisition in Interlanguage Pragmatics. Learning how to do things with words in a study abroad context. 2003. xviii, 403 pp.

107 **TAAVITSAINEN, Irma and Andreas H. JUCKER (eds.):** Diachronic Perspectives on Address Term Systems. 2003. viii, 446 pp.

106 **BUSSE, Ulrich:** Linguistic Variation in the Shakespeare Corpus. Morpho-syntactic variability of second person pronouns. 2002. xiv, 344 pp.

105 **BLACKWELL, Sarah E.:** Implicatures in Discourse. The case of Spanish NP anaphora. 2003. xvi, 303 pp.

104 **BEECHING, Kate:** Gender, Politeness and Pragmatic Particles in French. 2002. x, 251 pp.

103 **FETZER, Anita and Christiane MEIERKORD (eds.):** Rethinking Sequentiality. Linguistics meets conversational interaction. 2002. vi, 300 pp.

102 **LEAFGREN, John:** Degrees of Explicitness. Information structure and the packaging of Bulgarian subjects and objects. 2002. xii, 252 pp.

101 **LUKE, K. K. and Theodossia-Soula PAVLIDOU (eds.):** Telephone Calls. Unity and diversity in conversational structure across languages and cultures. 2002. x, 295 pp.

100 **JASZCZOLT, Katarzyna M. and Ken TURNER (eds.):** Meaning Through Language Contrast. Volume 2. 2003. viii, 496 pp.

99 **JASZCZOLT, Katarzyna M. and Ken TURNER (eds.):** Meaning Through Language Contrast. Volume 1. 2003. xii, 388 pp.

98 **DUSZAK, Anna (ed.):** Us and Others. Social identities across languages, discourses and cultures. 2002. viii, 522 pp.

97 **MAYNARD, Senko K.:** Linguistic Emotivity. Centrality of place, the topic-comment dynamic, and an ideology of *pathos* in Japanese discourse. 2002. xiv, 481 pp.

96 **HAVERKATE, Henk:** The Syntax, Semantics and Pragmatics of Spanish Mood. 2002. vi, 241 pp.

95 **FITZMAURICE, Susan M.:** The Familiar Letter in Early Modern English. A pragmatic approach. 2002. viii, 263 pp.

94 **McILVENNY, Paul (ed.):** Talking Gender and Sexuality. 2002. x, 332 pp.

93 **BARON, Bettina and Helga KOTTHOFF (eds.):** Gender in Interaction. Perspectives on femininity and masculinity in ethnography and discourse. 2002. xxiv, 357 pp.

92 **GARDNER, Rod:** When Listeners Talk. Response tokens and listener stance. 2001. xxii, 281 pp.

91 **GROSS, Joan:** Speaking in Other Voices. An ethnography of Walloon puppet theaters. 2001. xxviii, 341 pp.

90 **KENESEI, István and Robert M. HARNISH (eds.):** Perspectives on Semantics, Pragmatics, and Discourse. A Festschrift for Ferenc Kiefer. 2001. xxii, 352 pp.

89 **ITAKURA, Hiroko:** Conversational Dominance and Gender. A study of Japanese speakers in first and second language contexts. 2001. xviii, 231 pp.

88 **BAYRAKTAROĞLU, Arın and Maria SIFIANOU (eds.):** Linguistic Politeness Across Boundaries. The case of Greek and Turkish. 2001. xiv, 439 pp.

87 **MUSHIN, Ilana:** Evidentiality and Epistemological Stance. Narrative Retelling. 2001. xviii, 244 pp.

86 **IFANTIDOU, Elly:** Evidentials and Relevance. 2001. xii, 225 pp.

85 **COLLINS, Daniel E.:** Reanimated Voices. Speech reporting in a historical-pragmatic perspective. 2001. xx, 384 pp.

84 **ANDERSEN, Gisle:** Pragmatic Markers and Sociolinguistic Variation. A relevance-theoretic approach to the language of adolescents. 2001. ix, 352 pp.

83 **MÁRQUEZ REITER, Rosina:** Linguistic Politeness in Britain and Uruguay. A contrastive study of requests and apologies. 2000. xviii, 225 pp.

82 **KHALIL, Esam N.:** Grounding in English and Arabic News Discourse. 2000. x, 274 pp.

81 **DI LUZIO, Aldo, Susanne GÜNTHNER and Franca ORLETTI (eds.):** Culture in Communication. Analyses of intercultural situations. 2001. xvi, 341 pp.

80 **UNGERER, Friedrich (ed.):** English Media Texts – Past and Present. Language and textual structure. 2000. xiv, 286 pp.

79 **ANDERSEN, Gisle and Thorstein FRETHEIM (eds.):** Pragmatic Markers and Propositional Attitude. 2000. viii, 273 pp.

78 **SELL, Roger D.:** Literature as Communication. The foundations of mediating criticism. 2000. xiv, 348 pp.

77 **VANDERVEKEN, Daniel and Susumu KUBO (eds.):** Essays in Speech Act Theory. 2002. vi, 328 pp.

76 **MATSUI, Tomoko:** Bridging and Relevance. 2000. xii, 251 pp.

75 **PILKINGTON, Adrian:** Poetic Effects. A relevance theory perspective. 2000. xiv, 214 pp.

74 **TROSBORG, Anna (ed.):** Analysing Professional Genres. 2000. xvi, 256 pp.

73 **HESTER, Stephen K. and David FRANCIS (eds.):** Local Educational Order. Ethnomethodological studies of knowledge in action. 2000. viii, 326 pp.

72 **MARMARIDOU, Sophia S.A.:** Pragmatic Meaning and Cognition. 2000. xii, 322 pp.

71 **GÓMEZ GONZÁLEZ, María de los Ángeles:** The Theme–Topic Interface. Evidence from English. 2001. xxiv, 438 pp.

70 **SORJONEN, Marja-Leena:** Responding in Conversation. A study of response particles in Finnish. 2001. x, 330 pp.

69 **NOH, Eun-Ju:** Metarepresentation. A relevance-theory approach. 2000. xii, 242 pp.

68 **ARNOVICK, Leslie K.:** Diachronic Pragmatics. Seven case studies in English illocutionary development. 2000. xii, 196 pp.

67 **TAAVITSAINEN, Irma, Gunnel MELCHERS and Päivi PAHTA (eds.):** Writing in Nonstandard English. 2000. viii, 404 pp.

66 **JUCKER, Andreas H., Gerd FRITZ and Franz LEBSANFT (eds.):** Historical Dialogue Analysis. 1999. viii, 478 pp.

65 **COOREN, François:** The Organizing Property of Communication. 2000. xvi, 272 pp.

64 **SVENNEVIG, Jan:** Getting Acquainted in Conversation. A study of initial interactions. 2000. x, 384 pp.

63 **BUBLITZ, Wolfram, Uta LENK and Eija VENTOLA (eds.):** Coherence in Spoken and Written Discourse. How to create it and how to describe it. Selected papers from the International Workshop on Coherence, Augsburg, 24-27 April 1997. 1999. xiv, 300 pp.

62 **TZANNE, Angeliki:** Talking at Cross-Purposes. The dynamics of miscommunication. 2000. xiv, 263 pp.

61 **MILLS, Margaret H. (ed.):** Slavic Gender Linguistics. 1999. xviii, 251 pp.

60 **JACOBS, Geert:** Preformulating the News. An analysis of the metapragmatics of press releases. 1999. xviii, 428 pp.

59 **KAMIO, Akio and Ken-ichi TAKAMI (eds.):** Function and Structure. In honor of Susumu Kuno. 1999. x, 398 pp.

58 **ROUCHOTA, Villy and Andreas H. JUCKER (eds.):** Current Issues in Relevance Theory. 1998. xii, 368 pp.

57 **JUCKER, Andreas H. and Yael ZIV (eds.):** Discourse Markers. Descriptions and theory. 1998. x, 363 pp.

56 **TANAKA, Hiroko:** Turn-Taking in Japanese Conversation. A Study in Grammar and Interaction. 2000. xiv, 242 pp.

55 **ALLWOOD, Jens and Peter GÄRDENFORS (eds.):** Cognitive Semantics. Meaning and cognition. 1999. x, 201 pp.

54 **HYLAND, Ken:** Hedging in Scientific Research Articles. 1998. x, 308 pp.

53 **MOSEGAARD HANSEN, Maj-Britt:** The Function of Discourse Particles. A study with special reference to spoken standard French. 1998. xii, 418 pp.

52 **GILLIS, Steven and Annick DE HOUWER (eds.):** The Acquisition of Dutch. With a Preface by Catherine E. Snow. 1998. xvi, 444 pp.

51 **BOULIMA, Jamila:** Negotiated Interaction in Target Language Classroom Discourse. 1999. xiv, 338 pp.

50 **GRENOBLE, Lenore A.:** Deixis and Information Packaging in Russian Discourse. 1998. xviii, 338 pp.

49 **KURZON, Dennis:** Discourse of Silence. 1998. vi, 162 pp.

48 **KAMIO, Akio:** Territory of Information. 1997. xiv, 227 pp.

47 **CHESTERMAN, Andrew:** Contrastive Functional Analysis. 1998. viii, 230 pp.

46 **GEORGAKOPOULOU, Alexandra:** Narrative Performances. A study of Modern Greek storytelling. 1997. xvii, 282 pp.

45 **PALTRIDGE, Brian:** Genre, Frames and Writing in Research Settings. 1997. x, 192 pp.

44 **BARGIELA-CHIAPPINI, Francesca and Sandra J. HARRIS:** Managing Language. The discourse of corporate meetings. 1997. ix, 295 pp.

43 **JANSSEN, Theo and Wim van der WURFF (eds.):** Reported Speech. Forms and functions of the verb. 1996. x, 312 pp.

42 **KOTTHOFF, Helga and Ruth WODAK (eds.):** Communicating Gender in Context. 1997. xxvi, 424 pp.

41 **VENTOLA, Eija and Anna MAURANEN (eds.):** Academic Writing. Intercultural and textual issues. 1996. xiv, 258 pp.

40 **DIAMOND, Julie:** Status and Power in Verbal Interaction. A study of discourse in a close-knit social network. 1996. viii, 184 pp.

39 **HERRING, Susan C. (ed.):** Computer-Mediated Communication. Linguistic, social, and cross-cultural perspectives. 1996. viii, 326 pp.

38 **FRETHEIM, Thorstein and Jeanette K. GUNDEL (eds.):** Reference and Referent Accessibility. 1996. xii, 312 pp.

37 **CARSTON, Robyn and Seiji UCHIDA (eds.):** Relevance Theory. Applications and implications. 1998. x, 300 pp.

36 **CHILTON, Paul, Mikhail V. ILYIN and Jacob L. MEY (eds.):** Political Discourse in Transition in Europe 1989–1991. 1998. xi, 272 pp.

35 **JUCKER, Andreas H. (ed.):** Historical Pragmatics. Pragmatic developments in the history of English. 1995. xvi, 624 pp.

34 **BARBE, Katharina:** Irony in Context. 1995. x, 208 pp.

33 **GOOSSENS, Louis, Paul PAUWELS, Brygida RUDZKA-OSTYN, Anne-Marie SIMON-VANDENBERGEN and Johan VANPARYS:** By Word of Mouth. Metaphor, metonymy and linguistic action in a cognitive perspective. 1995. xii, 254 pp.

32 **SHIBATANI, Masayoshi and Sandra A. THOMPSON (eds.):** Essays in Semantics and Pragmatics. In honor of Charles J. Fillmore. 1996. x, 322 pp.

31 **WILDGEN, Wolfgang:** Process, Image, and Meaning. A realistic model of the meaning of sentences and narrative texts. 1994. xii, 281 pp.

30 **WORTHAM, Stanton E.F.:** Acting Out Participant Examples in the Classroom. 1994. xiv, 178 pp.

29 **BARSKY, Robert F.:** Constructing a Productive Other. Discourse theory and the Convention refugee hearing. 1994. x, 272 pp.

28 **VAN DE WALLE, Lieve:** Pragmatics and Classical Sanskrit. A pilot study in linguistic politeness. 1993. xii, 454 pp.

27 **SUTER, Hans-Jürg:** The Wedding Report. A prototypical approach to the study of traditional text types. 1993. xii, 314 pp.

26 **STYGALL, Gail:** Trial Language. Differential discourse processing and discursive formation. 1994. xii, 226 pp.

25 **COUPER-KUHLEN, Elizabeth:** English Speech Rhythm. Form and function in everyday verbal interaction. 1993. x, 346 pp.

24 **MAYNARD, Senko K.:** Discourse Modality. Subjectivity, Emotion and Voice in the Japanese Language. 1993. x, 315 pp.

23 **FORTESCUE, Michael, Peter HARDER and Lars KRISTOFFERSEN (eds.):** Layered Structure and Reference in a Functional Perspective. Papers from the Functional Grammar Conference, Copenhagen, 1990. 1992. xiii, 444 pp.

22 **AUER, Peter and Aldo DI LUZIO (eds.):** The Contextualization of Language. 1992. xvi, 402 pp.

21 **SEARLE, John R., Herman PARRET and Jef VERSCHUEREN:** (On) Searle on Conversation. Compiled and introduced by Herman Parret and Jef Verschueren. 1992. vi, 154 pp.

20 **NUYTS, Jan:** Aspects of a Cognitive-Pragmatic Theory of Language. On cognition, functionalism, and grammar. 1991. xii, 399 pp.

19 **BAKER, Carolyn and Allan LUKE (eds.):** Towards a Critical Sociology of Reading Pedagogy. Papers of the XII World Congress on Reading. 1991. xxi, 287 pp.

18 **JOHNSTONE, Barbara:** Repetition in Arabic Discourse. Paradigms, syntagms and the ecology of language. 1991. viii, 130 pp.

17 **PIÉRAUT-LE BONNIEC, Gilberte and Marlene DOLITSKY (eds.):** Language Bases ... Discourse Bases. Some aspects of contemporary French-language psycholinguistics research. 1991. vi, 342 pp.

16 **MANN, William C. and Sandra A. THOMPSON (eds.):** Discourse Description. Diverse linguistic analyses of a fund-raising text. 1992. xiii, 409 pp.

15 **KOMTER, Martha L.:** Conflict and Cooperation in Job Interviews. A study of talks, tasks and ideas. 1991. viii, 252 pp.

14 **SCHWARTZ, Ursula V.:** Young Children's Dyadic Pretend Play. A communication analysis of plot structure and plot generative strategies. 1991. vi, 151 pp.

13 **NUYTS, Jan, A. Machtelt BOLKESTEIN and Co VET (eds.):** Layers and Levels of Representation in Language Theory. A functional view. 1990. xii, 348 pp.

12 **ABRAHAM, Werner (ed.):** Discourse Particles. Descriptive and theoretical investigations on the logical, syntactic and pragmatic properties of discourse particles in German. 1991. viii, 338 pp.

11 **LUONG, Hy V.:** Discursive Practices and Linguistic Meanings. The Vietnamese system of person reference. 1990. x, 213 pp.

10 **MURRAY, Denise E.:** Conversation for Action. The computer terminal as medium of communication. 1991. xii, 176 pp.

9 **LUKE, K. K.:** Utterance Particles in Cantonese Conversation. 1990. xvi, 329 pp.

8 **YOUNG, Lynne:** Language as Behaviour, Language as Code. A study of academic English. 1991. ix, 304 pp.

7 **LINDENFELD, Jacqueline:** Speech and Sociability at French Urban Marketplaces. 1990. viii, 173 pp.

6:3 **BLOMMAERT, Jan and Jef VERSCHUEREN (eds.):** The Pragmatics of International and Intercultural Communication. Selected papers from the International Pragmatics Conference, Antwerp, August 1987. Volume 3: The Pragmatics of International and Intercultural Communication. 1991. viii, 249 pp.

6:2 **VERSCHUEREN, Jef (ed.):** Levels of Linguistic Adaptation. Selected papers from the International Pragmatics Conference, Antwerp, August 1987. Volume 2: Levels of Linguistic Adaptation. 1991. viii, 339 pp.

6:1 **VERSCHUEREN, Jef (ed.):** Pragmatics at Issue. Selected papers of the International Pragmatics Conference, Antwerp, August 17–22, 1987. Volume 1: Pragmatics at Issue. 1991. viii, 314 pp.

5 **THELIN, Nils B. (ed.):** Verbal Aspect in Discourse. 1990. xvi, 490 pp.

4 **RAFFLER-ENGEL, Walburga von (ed.):** Doctor–Patient Interaction. 1989. xxxviii, 294 pp.

3 **OLEKSY, Wieslaw (ed.):** Contrastive Pragmatics. 1988. xiv, 282 pp.

2 **BARTON, Ellen:** Nonsentential Constituents. A theory of grammatical structure and pragmatic interpretation. 1990. xviii, 247 pp.

1 **WALTER, Bettyruth:** The Jury Summation as Speech Genre. An ethnographic study of what it means to those who use it. 1988. xvii, 264 pp.