

Douglas Biber and Bethany Gray



Grammatical Complexity in Academic English

Linguistic Change in Writing

GRAMMATICAL COMPLEXITY IN ACADEMIC ENGLISH

Grammatical Complexity in Academic English uses corpus-based analyses to challenge a number of dominant stereotypes and assumptions within linguistics. Biber and Gray tackle the nature of grammatical complexity, demonstrating that embedded phrasal structures are as important as embedded dependent clauses. The authors also overturn ingrained assumptions about linguistic change, showing that grammatical change occurs in writing as well as speech. This work establishes that academic writing is structurally compressed (rather than elaborated); that it is often not explicit in the expression of meaning; and that scientific academic writing has been the locus of some of the most important grammatical changes in English over the past 200 years (rather than being conservative and resistant to change).

Supported throughout with textual evidence, this work is essential reading for discourse analysts, sociolinguists, and applied linguists, as well as descriptive linguists and historical linguists.

DOUGLAS BIBER is Regents' Professor of Applied Linguistics in the English Department at Northern Arizona University.

BETHANY GRAY is Assistant Professor of English (Applied Linguistics and Technology) at Iowa State University.

STUDIES IN ENGLISH LANGUAGE

General editor: Merja Kytö (Uppsala University)

Editorial board: Bas Aarts (University College London), John Algeo (University of Georgia), Susan Fitzmaurice (University of Sheffield), Christian Mair (University of Freiburg), Charles F. Meyer (University of Massachusetts)

The aim of this series is to provide a framework for original studies of English, both present day and past. All books are based securely on empirical research and represent theoretical and descriptive contributions to our knowledge of national and international varieties of English, both written and spoken. The series covers a broad range of topics and approaches, including syntax, phonology, grammar, vocabulary, discourse, pragmatics and sociolinguistics, and is aimed at an international readership.

Already published in this series

JIM FEIST:

Premodifiers in English: Their Structure and Significance

STEVEN JONES, M. LYNNE MURPHY, CARITA PARADIS, AND CAROLINE
WILLNERS:

Antonyms in English: Construals, Constructions and Canonicity

CHRISTIANE MEIERKORD:

Interactions across Englishes: Linguistic Choices in Local and International
Contact Situations

HARUKO MOMMA:

From Philology to English Studies: Language and Culture in the Nineteenth
Century

RAYMOND HICKEY (ED.):

Standards of English: Codified Varieties Around the World

BENEDIKT SZMRECSANYI:

Grammatical Variation in British English Dialects: A Study in Corpus-Based
Dialectometry

DANIEL SCHREIER AND MARIANNE HUNDT (EDS.):

English as a Contact Language

BAS AARTS, JOANNE CLOSE, GEOFFREY LEECH, AND SEAN WALLIS (EDS.):

The Verb Phrase in English: Investigating Recent Language Change with Corpora

MARTIN HILPERT:

Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax

JAKOB R. E. LEIMGRUBER:

Singapore English: Structure, Variation and Usage

CHRISTOPH RÜHLEMANN:

Narrative in English Conversation

DAGMAR DEUBER:

English in the Caribbean: Variation, Style and Standards in Jamaica and Trinidad

EVA BERLAGE:

Noun Phrase Complexity in English

NICOLE DEHÉ:

Parentheticals in Spoken English: The Syntax-Prosody Relation

JOCK ONN WONG:

English in Singapore: A Cultural Analysis

ANITA AUER, DANIEL SCHREIER, AND RICHARD J. WATTS:

Letter Writing and Language Change

MARIANNE HUNDT:

Late Modern English Syntax

IRMA TAAVITSAINEN, MERJA KYTO, CLAUDIA CLARIDGE, AND JEREMY SMITH:

Developments in English: Expanding Electronic Evidence

ARNE LOHMANN:

English Co-ordinate Constructions: A Processing Perspective on Constituent Order

JOHN FLOWERDEW AND RICHARD W. FOREST:

Signalling Nouns in English: A Corpus-Based Discourse Approach

JEFFREY P. WILLIAMS, EDGAR W. SCHNEIDER, PETER TRUDGILL, AND

DANIEL SCHREIER:

Further Studies in the Lesser-Known Varieties of English

NURIA YÁÑEZ-BOUZA:

Grammar, Rhetoric and Usage in English: Preposition Placement 1500–1900

JACK GRIEVE:

Regional Variation in Written American English

DOUGLAS BIBER AND BETHANY GRAY:

Grammatical Complexity in Academic English: Linguistic Change in Writing

Earlier titles not listed are also available

GRAMMATICAL COMPLEXITY IN ACADEMIC ENGLISH

Linguistic Change in Writing

DOUGLAS BIBER

Northern Arizona University

BETHANY GRAY

Iowa State University



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107009264

© Douglas Biber and Bethany Gray 2016

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2016

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication Data

Biber, Douglas, author.

Grammatical complexity in academic English : linguistic change in writing / Douglas Biber, Northern Arizona University ; Bethany Gray, Iowa State University.

pages cm. – (Studies in English language)

Includes bibliographical references and index.

ISBN 978-1-107-00926-4 (Hardback)

1. English language—Grammar. 2. Complexity (Linguistics) 3. Academic writing—Research.
4. Linguistic change. I. Gray, Bethany, author. II. Title.

P128.C664B47 2015

425—dc23 2015028909

ISBN 978-1-107-00926-4 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

<i>List of figures</i>	<i>page</i> x
<i>List of tables</i>	xiii
1 Academic writing: Challenging the stereotypes 1	
1.1 Academese: Obtuse or informational? 1	
1.2 Describing the grammar of academic writing 5	
1.2.1 Academic written texts: All basically the same? 7	
1.2.2 Academic writing: Complex grammar and explicit meanings? 14	
1.2.3 Academic writing: Resistant to change or dynamic? 19	
1.3 Grammatical change in English: Below the radar? 27	
1.4 Two types of grammatical change 30	
1.5 The locus of historical change in English: Speech or writing? 32	
1.6 Overview of the present book 39	
2 Using corpora to analyze grammatical change 43	
2.1 Introduction 43	
2.2 Situational change within academic research writing 50	
2.3 Corpora analyzed in the book 51	
2.3.1 Representing academic writing over time: Description of the diachronic academic corpora 52	
2.3.2 Synchronic and diachronic comparisons between academic writing and other spoken and written registers 57	
2.4 Corpus analysis procedures 58	
2.5 Grammatical features investigated 59	
2.5.1 'Core' grammatical features and 'colloquial' features 60	
2.5.2 Phrasal and clausal complexity features 60	
2.5.3 Corpus analyses of grammatical complexity features 65	
3 Phrasal versus clausal discourse styles: A synchronic grammatical description of academic writing contrasted with other registers 67	
3.1 Introduction 67	
3.2 Previous linguistic research on academic writing 72	

3.2.1	Studies of particular grammatical features in academic writing	74
3.2.2	A comprehensive survey of the distinctive grammatical features of academic writing: Contributions from <i>The Longman Grammar of Spoken and Written English</i>	76
3.2.3	Studies of register variation that include academic writing	83
3.3	Case studies illustrating the distinctive grammatical characteristics of academic research writing	87
3.3.1	Academic writing contrasted with conversation: Which is more grammatically complex?	87
3.3.2	University textbooks contrasted with university classroom teaching: Is there a general academic style?	100
3.3.3	Grammatical features of academic writing contrasted with popular written registers	104
3.3.4	Grammatical variation among written sub-registers from different academic disciplines	111
3.4	Chapter summary	123
4	The historical evolution of phrasal discourse styles in academic writing	125
4.1	Grammatical complexity in eighteenth century written texts	125
4.1.1	The influence of popularization versus economy in historical change	128
4.1.2	Previous historical research on the noun phrase in English	130
4.2	General patterns of change in written English registers	132
4.2.1	Increasing use of colloquial features	136
4.2.2	Historical change in the use of complexity features	140
4.2.3	Exploring the magnitude of change: The evolution of a phrasal discourse style	145
4.3	What features have decreased in use in academic prose?	154
4.4	Historical change across sub-registers within academic research writing	157
4.5	Chapter summary	166
5	The functional extension of phrasal grammatical features in academic writing	167
5.1	Introduction	167
5.2	Phrasal features functioning as pre-modifiers of a head noun	170
5.2.1	Nouns as nominal pre-modifiers	170
5.2.2	Attributive adjectives as nominal pre-modifiers	184
5.2.3	Noun-participle compounds as nominal pre-modifiers	187
5.3	Phrasal features functioning as post-modifiers of a head noun	190
5.3.1	Prepositional phrases as nominal post-modifiers	190
5.3.2	Appositive noun phrases as nominal post-modifiers	202
5.4	The systemic 'drift' towards structural compression, motivated by economy of expression	207

6	The loss of explicitness in academic research writing	218
6.1	Stereotypes concerning explicitness in academic research writing	218
6.2	Explicitness of reference in conversation versus academic writing	219
6.3	Inexplicit meaning relationships associated with phrasal pre-modifiers in noun phrases	222
6.4	Inexplicit meaning relationships associated with phrasal post-modifiers in noun phrases	232
6.5	Inexplicit expression associated with compressed clausal connectors in academic writing	240
7	Conclusion	244
7.1	Stereotypes about grammatical complexity: What does it mean to be 'complex'?	245
7.2	Stereotypes about historical change: Does all grammatical change originate in speech?	246
7.3	Stereotypes about academic writing: Is it resistant to change, elaborated, and explicit?	248
7.4	Implications for applied linguists: From language development to language teaching and assessment	251
7.5	Conclusion	256
	<i>References</i>	257
	<i>Appendix 1 Descriptive statistics for nine linguistic features for three general registers across historical periods</i>	272
	<i>Appendix 2 Descriptive statistics for five linguistic features for sub-registers of academic writing across historical periods</i>	275
	<i>Index</i>	276

Figures

2.1	Decrease in the use of modal verbs in AmE fiction and newspapers (based on COHA)	<i>page 45</i>
2.2	Increase in the use of the semi-modal <i>have to</i> in AmE fiction and newspapers (based on COHA)	46
2.3	Increase in the use of specific phrasal-prepositional verbs in AmE (based on COHA)	47
2.4	A cline of grammatical complexity	62
3.1	Academic and non-academic registers along two situational parameters	69
3.2	The distribution of complexity features in conversation and academic writing	92
3.3	Common finite clause types functioning as clausal constituents	95
3.4	Common dependent phrasal types functioning as constituents in a noun phrase	96
3.5	Dependent structures that mix the two parameters	98
3.6	Major grammatical classes: academic speech and writing	102
3.7	Finite dependent clauses: academic speech and writing	103
3.8	Major grammatical classes across three major written registers	106
3.9	Dependent clause types across three major written registers	108
3.10	Noun phrase modifiers across three major written registers	110
3.11	Major grammatical classes across sub-disciplines of academic writing	114
3.12	Dependent clause types across sub-disciplines of academic writing	115
3.13	Noun phrase modifiers across sub-disciplines of academic writing	116
3.14	Prepositional phrases across sub-disciplines of academic writing	118

3.15	Appositive noun phrases in newspapers versus two sub-disciplines of academic writing	119
3.16	Noun + participle as nominal pre-modifiers across sub-disciplines of academic writing	121
3.17	Explicit versus implicit signals of logical relations across sub-disciplines of academic writing	122
4.1	Distribution of nouns (per 1,000 words) across historical periods	134
4.2	Historical change in the use of progressive aspect verbs	138
4.3	Historical change in the use of nouns	146
4.4	Historical change in science research writing: common nouns, nominalizations, and verbs	147
4.5	Historical change in the use of noun phrase pre-modifier types in academic prose: attributive adjectives versus nouns as pre-modifiers	148
4.6	Historical change in the use of emerging NP pre-modifier types in academic prose: N-XXing + N and N-XXed + N	149
4.7	Historical change in the use of finite relative clauses	150
4.8	Historical change in the distribution and functions of prepositional phrases in science academic prose	151
4.9	Historical change in the distribution of specific prepositional phrases functioning as noun modifiers in science academic prose	152
4.10	Appositive noun phrases as nominal post-modifiers	153
4.11	Historical change in the use of definite and indefinite articles in science academic prose	155
4.12	Historical change in the use of nouns across academic sub-registers	160
4.13	Historical change in the use of nominalizations across academic sub-registers	162
4.14	Historical change in the use of relative clauses across academic sub-registers	163
4.15	Historical change in the use of noun + <i>of</i> -phrase across academic sub-registers	164
4.16	Historical change in the use of nouns as nominal pre-modifiers across academic sub-registers	165
5.1	Historical change in the proportional use of pre-modifying noun genitives (versus <i>'s</i> -genitives)	173
5.2	Historical change in the proportional use of pre-modifying noun genitives (versus <i>of</i> -genitives)	174

5.3	The 100 most frequent adjective types, categorized for descriptive versus classifier functions	185
5.4	IN as noun-modifier: concrete versus abstract meanings	193
5.5	ON as noun-modifier: concrete versus abstract meanings	194
5.6	Noun + preposition + <i>ing</i> -clause in academic prose	200
5.7	Historical change in the use of predicative relative clauses in science research articles	209
5.8	Historical change (in COHA) of three grammatical variants: N + of, N + on, N + N	213
5.9	Historical change (in COHA) of three grammatical variants: N + of, N + in, N + N	215
5.10	Historical change (based on COCA) of three grammatical variants from 1990–2012: N + of, N + in, N + N	216
6.1	Historical change in the use of linking adverbials versus colons as clause connectors in academic prose	242

Tables

1.1 Preferred grammatical features of Text Sample 1.3 (literary criticism) versus Text Sample 1.4 (science)	<i>page</i> 11
1.2 Characterizations of academic writing as elaborated and/or explicit	15
1.3 Summary of major grammatical differences between four text excerpts	26
1.4 Challenging basic assumptions about grammatical complexity, linguistic change, and academic writing through corpus-based analyses	41
2.1 Academic research journals sampled for the 20 th Century <i>Research Article Corpus</i>	53
2.2 Design and composition of the 20 th Century <i>Research Article Corpus</i>	54
2.3 Complete academic writing corpus used for long-term historical analyses in the book	55
2.4 Description of comparison corpora	58
2.5 Colloquial grammatical features investigated in the study	61
2.6 Example complexity features, by structural type versus syntactic function	61
2.7 Complexity features investigated in the study, by structural type versus syntactic function	63
3.1 Grammatical features that are especially common in academic prose (based on a survey of the <i>Longman Grammar of Spoken and Written English</i>)	79
3.2 The oral/literate dimension (Factor 1) in selected MD studies of particular discourse domains in English	85
3.3 Statistical comparisons for finite dependent clause types	93
3.4 Statistical comparisons for non-finite dependent clause types	93

3.5	Statistical comparisons for dependent phrase types (non-clausal)	94
3.6	Statistical comparisons for major word classes across three general written registers	105
3.7	Statistical comparisons for dependent clause features across three general written registers	107
3.8	Statistical comparisons for phrasal modifiers across three general written registers	109
3.9	Statistical comparisons for major word classes across disciplines in academic writing	111
3.10	Statistical comparisons for dependent clause features across disciplines in academic writing	112
3.11	Statistical comparisons for phrasal noun modifiers across disciplines in academic writing	113
4.1	Historical change in the use of general linguistic features, shown by Pearson correlation coefficients (r) for the rate of occurrence correlated with date (1750–1990)	135
4.2	Historical change in the use of colloquial features, shown by Pearson correlation coefficients for the rate of occurrence correlated with date (1800–1990)	137
4.3	Historical change in the use of complexity features, shown by Pearson correlation coefficients (r) for the rate of occurrence correlated with date (1750–1990)	142
4.4	Summary of the ANOVA factorial models for four academic sub-registers (specialist science, specialist social science, multi-disciplinary science, humanities) in three recent historical periods (1965, 1985, 2005)	161
6.1	Statistical comparisons for clause connectors signaling logical relations	241
7.1	Hypothesized developmental stages for complexity features	253
A1.1	Descriptive statistics for lexical verbs, adjectives, and adverbs	272
A1.2	Descriptive statistics for passive voice verbs, nominalizations, and prepositional phrases	273
A1.3	Descriptive statistics for finite adverbial clauses, complement clauses, and relative clauses	274
A2.1	Descriptive statistics for five grammatical features in academic sub-registers from three recent historical periods (1965, 1985, 2005) (rates of occurrence per 1,000 words)	275

Academic writing: Challenging the stereotypes

1.1 Academese: Obtuse or informational?

Many speakers of English share the view that the language of academic writing is peculiar, not only different from everyday speech, but also different from most other registers of English. These perceived differences are not neutral. Rather, a common stereotype of academic prose is that it is deliberately complex, and more concerned with impressing readers than communicating ideas – all making it needlessly difficult to understand.

These attitudes are reflected in the label *academese*, which is usually defined with negative connotations. For example, the *Tameri Guide for Writers* (www.tameri.com) defines *academese* as ‘an artificial form of communication commonly used in institutes of higher education designed to make small, irrelevant ideas appear important and original. Proficiency in academese is achieved when you begin inventing your own words and no one can understand what you are writing.’

In fact, the suffix *-ese* seems to have taken on this broader meaning when attached to other roots. Thus, according to the on-line site www.wiktionary.org, *officialese* is ‘the typical language of [written] official documents, legalistic and pompous’; *legalese* is ‘wording that resembles how a lawyer writes, especially such that is confusing to the layperson’; and *bureaucratese* is ‘any language containing many non-essential words intended to imply more importance or intelligence than actually present’.

The primary focus of these negative attitudes is on the use of rare and obscure words and phrases. The perception is that these words are often chosen to impress readers rather than efficiently convey new information. The portrayal of academic prose in similar terms has been around for some time, as evidenced by Orwell’s famous critique in 1946:

The writer either has a meaning and cannot express it, or he inadvertently says something else, or he is almost indifferent as to whether

his words mean anything or not. This mixture of vagueness and sheer incompetence is the most marked characteristic of modern English prose. . .

In certain kinds of writing, particularly in art criticism and literary criticism, it is normal to come across long passages which are almost completely lacking in meaning.

Orwell, George. 1946.
'Politics and the English Language'
Horizon, 13(76): 252–265

Here we see Orwell being especially critical of academic writing in the humanities (specifically art criticism and literary criticism). The enterprise of research in the humanities is focused on our everyday experience. Unlike scientists, humanities researchers are not discovering and documenting new natural phenomena and processes. Rather, their work is more interpretive, attempting to describe and understand the human experience. Like the author of Ecclesiastes, readers might believe that 'there is nothing new under the sun' when it comes to discussions of the human experience – and as a result, expect that there would be little need for new technical vocabulary and complex grammatical constructions to discuss these familiar ideas, relationships, and experiences. This seems to be Orwell's underlying complaint: humanities researchers mostly discuss phenomena that are already familiar to us all, but they do it in a style that is indecipherable to non-specialist readers.

These stereotypes about academese in humanities writing continue right up to the present. For example, Pinker (2014: para 4) discusses the question of 'Why academics stink at writing?', and notes that:

The most popular answer outside the academy is the cynical one: [...] Scholars in the softer fields spout obscure verbiage to hide the fact that they have nothing to say. They dress up the trivial and obvious with the trappings of scientific sophistication, hoping to bamboozle their audiences with highfalutin gobbledygook.

Texts that seem to illustrate this prose style are unfortunately not hard to find. For example, the journal *Philosophy and Literature* sponsored a 'Bad Writing Contest' from 1995–1998, which celebrated 'the most stylistically lamentable passages found in scholarly books and articles published in the last few years' (http://denisdutton.com/bad_writing.htm). The contest lampooned academese simply by quoting the academic prose of famous literary scholars. For example, the winner of the 1998 contest was Judith Butler, 'a Guggenheim Fellowship-winning professor of rhetoric and comparative literature at the University of California at Berkeley'. The first-prize sentence, singled out for its 'anxiety-inducing obscurity', reads as follows:

Text Sample 1.1
Literary criticism article

The move from a structuralist account in which capital is understood to structure social relations in relatively homologous ways to a view of hegemony in which power relations are subject to repetition, convergence, and rearticulation brought the question of temporality into the thinking of structure, and marked a shift from a form of Althusserian theory that takes structural totalities as theoretical objects to one in which the insights into the contingent possibility of structure inaugurate a renewed conception of hegemony as bound up with the contingent sites and strategies of the rearticulation of power.

Judith Butler. 1997.
'Further Reflections on the Conversations of Our Time'
Diacritics, 27(1): 13–15.

As noted earlier, the popular explanation for academese in humanities writing is that it represents a form of deliberate obscurity. Pinker (2014), however, proposes a slightly different explanation, based on the underlying goals of humanities versus science research writing. Pinker describes the primary goal of humanities writing as 'self-presentation' rather than the communication of information. This goal results in a 'self-conscious style', where the author's primary concern is 'to escape being convicted of philosophical naïveté about his own enterprise'. In contrast, science research writing is associated with a 'classic style' focused on the communication of information: 'The writer can see something that the reader has not yet noticed, and he orients the reader so she can see for herself. The purpose of writing is presentation, and its motive is disinterested truth. [. . .] The writer knows the truth before putting it into words; he is not using the occasion of writing to sort out what he thinks' (Pinker 2014: para 11).

The linguistic characteristics of the 'classic' style associated with science writing are dramatically different from the 'self-conscious' style of humanities writing. Science prose can be boring, relying on simple syntactic constructions, as in:

Text Sample 1.2
Biology research article

The neurites are black on a yellow-brown background in the original preparation (see Figure 5). One neurite can be traced coursing through the basement membrane of the epidermis (arrow). The neurites appear to penetrate the cytoplasm of the epidermal cells (see also Fig. 11).

Bryce L. Munger. 1965.
'The Intraepidermal Innervation of the Snout Skin of the Opossum'
Journal of Cell Biology, 26(1): 79–97.

Although there are often technical terms in this style, those terms usually refer to physical entities rather than abstract concepts. The overall stylistic effect is to emphasize the direct communication of information (even if that information is still unclear to non-specialist readers because of the technical vocabulary).

Our goals in the present book are not to enter into the debate about ‘good’ and ‘bad’ academic writing, or the motivations for the ‘academese’ found in many academic texts. But we *are* interested in the linguistic characteristics of academic writing, including the differences between humanities writing and science writing. And we are especially interested in the ways in which academic writing has changed historically in its grammatical style.

Given the striking linguistic differences between present-day humanities writing versus science writing, readers probably will not be surprised to learn that academic writing styles have undergone major linguistic change over the past few centuries. In particular, it would be easy to assume that academic writing has become increasingly complex over time, and that humanities writing has been especially influenced by these changes, resulting in the grammatically-elaborated styles of academese illustrated in [Text Sample 1.1](#). At the same time, it might be assumed that science prose has somehow resisted these historical changes, and thus retained the grammatically-simpler styles of earlier centuries.

It turns out that those expectations are wrong in nearly every way. In the following chapters, we show that:

- There are different types of grammatical complexity.
- Complexity in humanities writing is associated with structural elaboration.
- Humanities academic writing has changed little over the last few centuries in its structurally elaborated grammatical style.
- Complexity in science writing is associated with structural compression, not structural elaboration.
- Science writing has undergone extreme historical change related to the use of these structural compression features.

In the next section, we provide more detailed discussion of the grammatical features associated with academic writing, and the striking grammatical differences between humanities and science writing. This is followed by sections that set the stage for our analyses: previous research on grammatical change, and the discourse factors that influence change. In contrast to most previous research – which focuses on spoken interaction as the breeding ground for linguistic innovations – we argue here that historical

change can occur in writing as well as speech. The following chapters present the results of empirical corpus analyses, showing that this theoretical possibility has indeed occurred. In fact, the corpus analyses in following chapters show that academic science writing has been the locus of some of the most dramatic grammatical changes that have occurred in English over the past three centuries.

1.2 Describing the grammar of academic writing

Researchers in the humanities would argue that they do not simply document human experiences that are familiar to us all; rather, they are constantly offering new interpretations of those experiences, and as a result, coining new terms that reflect those new interpretations. Thus, although these researchers discuss familiar situations and experiences, they do so in highly technical ways that are not easily understood by the non-expert. For example, consider the following excerpt from a literary criticism article:

Text Sample 1.3

Modern literary criticism article

THE MANUSCRIPT OF CHARLOTTE BRONTE'S "Villette" (1853)

Published in the wake of the Great Exhibition of 1851, it maps out the contours of interiority in a world newly captivated by the peculiar resonance of things. Though Bronte liked to think that her novel "touche[d] on no matter of public interest," its conception of the psychological interior was significantly inflected by its setting in mid century Thing City (Letters 3: 75). Villette places interiority in an intimate connection with object-filled interiors even as it hopes for an inner life that eludes the varied fetishisms of Thing City. This nostalgia for a more pristine and private form of psychological depth is, in turn, articulated in terms that reveal how entrenched persons are in the public empire of things. Villette constitutes an attempt to negotiate between a critique of commodity fetishism and a paradoxically fetishistic preoccupation with the traces and tokens of inner life. The novel suggests that the bourgeois subject, though it comes into being through its relations with things, is defined by the nostalgic notion that its true interiority has been lost under the pressure of things.

Eva Badowska. 2005.

'Choseville: Brontë's "Villette" and the Art of Bourgeois Interiority'

PMLA, 120: 1509–1523.

One of the most salient characteristics of this text is the highly specialized vocabulary. Abstract technical terms are common in this passage, such as *interiority* and *fetishisms*. In addition, relatively common words like *things*,

objects, contours, interior, depth, and life are used together with technical terms to produce complex phrases referring to highly abstract concepts, such as *the peculiar resonance of things, object-filled interiors, the contours of interiority, nostalgia for a more pristine and private form of psychological depth, the public empire of things, and a paradoxically fetishistic preoccupation with the traces and tokens of inner life*. Taken together, these abstract terms and complex phrases make it difficult for the non-specialist reader to understand the content of the literary analysis presented in the article.

Research writing in the sciences is probably even more difficult for the non-expert to understand, but its linguistic style is less often singled out for criticism. In part, this is because there is less of a mismatch between our expectations and the discourse style. The general goal of science research is to discover new information about the natural world, identifying new phenomena and explaining previously identified phenomena and patterns. Scientists require new words to refer to these previously undocumented phenomena and processes, resulting in prose that can be almost unintelligible to the non-expert. However, since the non-specialist reader has no prior experience with or existing knowledge of those phenomena and processes, we have little expectation that we should fully understand these texts. The following passage from a biochemistry research article illustrates:

Text Sample 1.4 Modern biochemistry article

Several kinases phosphorylate vimentin, the most common intermediate filament protein, in mitosis. Aurora-B and Rho-kinase regulate vimentin filament separation through the cleavage furrow-specific vimentin phosphorylation. Cdk1 also phosphorylates vimentin from prometaphase to metaphase, but its significance has remained unknown. Here we demonstrated a direct interaction between Plk1 and vimentin-Ser55 phosphorylated by Cdk1, an event that led to Plk1 activation and further vimentin phosphorylation. Plk1 induced the phosphorylation of vimentin-Ser82, which was elevated from metaphase and maintained until the end of mitosis. This elevation followed the Cdk1-induced vimentin-Ser55 phosphorylation, and was impaired by Plk1 depletion.

T. Yamaguchi et al. 2005.
'Phosphorylation by Cdk1 induces Plk1-mediated
vimentin phosphorylation during mitosis'
Journal of Cell Biology, 171(3): 431–436.

Similar to [Text Sample 1.3](#), the most obvious characteristic of this biochemistry article is its technical vocabulary. Most readers of the present book have never encountered the verb *to phosphorylate*, or nouns like *kinases, vimentin, prometaphase, metaphase, and phosphorylation*. Since we

have no idea what these terms refer to, the entire passage has little meaning for us. Thus, the vocabulary of science research writing is similarly technical to the vocabulary of humanities prose, supporting the stereotype that all academic writing is complex and hard to understand.

It is much more difficult to notice the typical grammatical structures used in these academic texts. However, when considered at the grammatical level, we discover that there are important linguistic differences among the various disciplines of academic prose; and in particular, the language of science research writing is quite different from the language of humanities prose. Such considerations lead in turn to other surprising findings that contradict previous claims and assumptions about academic writing. In particular, the corpus-based analyses presented in the following chapters challenge four major stereotypes about academic prose:

1. all kinds of academic prose are essentially the same
2. academic prose employs complex and elaborated grammar
3. academic prose is maximally explicit in meaning
4. academic prose is conservative and resistant to linguistic change

Along the way, we challenge two basic theoretical assumptions that have been widely adopted in previous linguistic research:

1. grammatical complexity is equivalent to structural elaboration, realized especially through the increased use of dependent clauses
2. grammatical changes are initiated in speech; grammatical innovations do not occur in writing

In the following sections, we briefly introduce each of these widely held beliefs and begin to challenge these assumptions.

1.2.1 *Academic written texts: All basically the same?*

Although writing researchers and students talk about ‘academic prose’ as if it were a well-defined construct, there are actually many ways in which academic texts differ from one another. For example, we discussed in preceding sections how humanities research writing differs in its goals and topics from science research writing. Despite those differences, though, all academic written texts can seem similar linguistically, especially in contrast to the discourse styles that we are familiar with in conversation or popular written registers like fiction.

Characteristics of this general academic-prose style are evident in both [Text Samples 1.3](#) and [1.4](#) above. For example, both texts rely on specialized technical vocabulary, including many nominalizations (nouns that are formed from verbs or adjectives by the addition of a suffix):

Literary criticism text:

*exhibition, interiority, resonance, conception, connection, fetishisms,
preoccupation, relations*

Biochemistry text:

*separation, cleavage, phosphorylation, significance, interaction, activation,
ability, elevation, depletion*

The two texts also share a reliance on some specialized grammatical features, which contribute to the perception that the texts are somehow peculiar and difficult to understand. For example, passive voice – a grammatical feature often associated with academic prose – is commonly used in both of these texts:

Literary criticism text:

a world newly captivated by ... things

its conception ... was significantly inflected by its setting

this nostalgia ... is, in turn, articulated ...

*the bourgeois subject ... is defined by the nostalgic notion that its true interiority
has been lost*

Biochemistry text:

vimentin-Ser⁵⁵ phosphorylated by CdkI

*the phosphorylation of vimentin-Ser⁸² ... was elevated from metaphase and
maintained until the end of mitosis*

This elevation ... was impaired by PlkI depletion

However, a more careful analysis uncovers ways in which these two texts differ in their grammatical characteristics, reflecting the different norms of their academic disciplines. For example, attributive adjectives (i.e., adjectives that pre-modify a head noun) are very common in humanities academic writing, and [Text Sample 1.3](#) above illustrates this characteristic:

Literary criticism text:

Great Exhibition

peculiar resonance

public interest

psychological interior

intimate connection

inner life

varied fetichisms

a more pristine and private form of psychological depth

public empire

a paradoxically fetichistic preoccupation

nostalgic notion

Attributive adjectives are generally less common in science research writing, and Text Sample 1.4 is typical in this regard, with only two examples: *direct interaction* and *further vimentin phosphorylation*.

However, science prose tends to employ a more specialized grammatical device to modify noun phrases: nouns that occur as pre-modifiers of a head noun. Here again, Text Sample 1.4 illustrates the typical pattern:

Biochemistry text:

filament protein

vimentin filament separation

the cleavage furrow-specific vimentin phosphorylation

Plk1 activation

vimentin phosphorylation

Plk1 depletion

These pre-modifying nouns can sometimes be compounded with participles, as in:

Cdk1-induced vimentin-Ser55 phosphorylation

In contrast, pre-modifying nouns are much less common in humanities writing, as is illustrated by the literary criticism text, which employs only two occurrences:

Literary criticism text:

Thing City

commodity fetishism

Further consideration of [Text Sample 1.4](#) illustrates an additional grammatical device that is much more common in science writing than humanities writing: appositive noun phrases. These are noun phrases that modify the immediately preceding head noun, with no overt grammatical connector. For example:

Biochemistry text:

vimentin, the most common intermediate filament protein

a direct **interaction** between *Plk1* and vimentin-Ser55 phosphorylated by Cdk1, an event that led to *Plk1* activation and further vimentin phosphorylation

[Table 1.1](#) summarizes these differences in the preferred grammatical features of the humanities versus science writing, illustrated from [Text Samples 1.3](#) and [1.4](#).

Comparing the grammatical structure of these two texts from a quantitative perspective uncovers additional differences. For example, the two passages have important differences in their preferred sentence structures. [Texts Samples 1.3a](#) and [1.4a](#) below are copies of the two text samples, highlighting the verbs in both passages. Although the literary criticism text is considerably longer than the biochemistry text (179 words versus 94 words), both passages consist of six sentences. Thus, the sentences are much longer in the literary criticism text than the biochemistry text. A related difference is the density of verbs in the two passages: the literary criticism text utilizes three to four verbs in each sentence, while the biochemistry text uses only one to two verbs per sentence.

Text Sample 1.3a

Literary criticism text (Badowska 2005)

Verbs marked in **bold underline**

Published in the wake of the Great Exhibition of 1851, it **maps** out the contours of interiority in a world newly **captivated** by the peculiar resonance of things. Though Bronte **liked** to **think** that her novel "**touche[d]** on no matter of public interest," its conception of the psychological interior **was** significantly **inflected** by its setting in mid century Thing City (Letters 3: 75). Villette **places** interiority in an intimate connection with object-filled interiors even as it **hopes** for an inner life that **eludes** the varied fetishisms of Thing City. This nostalgia for a more pristine and private form of psychological depth **is**, in turn, **articulated** in terms that **reveal** how **entrenched** persons **are** in

Table 1.1 *Preferred grammatical features of Text Sample 1.3 (literary criticism) versus Text Sample 1.4 (science)*

Linguistic Characteristic	Pattern	Literary Criticism Text (Text Sample 1.3)	Biochemistry Text (Text Sample 1.4)
Technical vocabulary and nominalizations	+ literary criticism + biochemistry	<i>exhibition, interiority, resonance, conception, connection, fetishisms, preoccupation, relations</i>	<i>separation, cleavage, phosphorylation, significance, interaction, activation, ability, elevation, depletion</i>
Passive voice	+ literary criticism + biochemistry	<i>a world newly <u>captivated</u> by ... things</i> <i>its conception ... <u>was significantly inflected</u> by its setting</i> <i>this nostalgia ... <u>is</u>, in turn, <u>articulated</u> ...</i> <i>the bourgeois subject ... <u>is defined</u> by the nostalgic notion that its true interiority <u>has been lost</u></i>	<i>vimentin-Ser55 <u>phosphorylated</u> by Cdkl</i> <i>the phosphorylation of vimentin-Ser82 ... <u>was elevated</u> from metaphase and <u>maintained</u> until the end of mitosis</i> <i>This elevation ... <u>was impaired</u> by Plk1 depletion</i>
Attributive Adjectives	+ literary criticism - biochemistry	<i><u>Great Exhibition</u>, <u>peculiar</u> resonance, <u>public</u> interest, <u>psychological</u> interior, <u>intimate</u> connection, <u>inner</u> life, <u>varied</u> fetishisms, a more <u>pristine</u> and <u>private</u> form of <u>psychological</u> depth, <u>public</u> empire, a paradoxically <u>fetishistic</u> preoccupation, <u>nostalgic</u> notion</i>	<i><u>direct</u> interaction, <u>further</u> vimentin phosphorylation</i>
Nouns as noun pre-modifiers	- literary criticism + biochemistry	<i><u>Thing</u> City, <u>commodity</u> fetishism</i>	<i><u>filament</u> protein, <u>vimentin</u> filament separation, the <u>cleavage furrow-specific</u> vimentin phosphorylation, <u>Plk1</u> activation, <u>vimentin</u> phosphorylation, <u>Plk1</u> depletion</i>
Noun + participle as noun pre-modifier	- literary criticism + biochemistry	– no examples –	<i><u>Cdkl-induced</u> vimentin-Ser55 phosphorylation</i>
Appositive noun phrases	- literary criticism + biochemistry	– no examples –	<i>vimentin, <u>the most common intermediate filament protein</u></i> <i>a direct interaction between Plk1 and vimentin-Ser55 phosphorylated by Cdkl, <u>an event that led to Plk1 activation and further vimentin phosphorylation</u></i>

the public empire of things. Villette **constitutes** an attempt to **negotiate** between a critique of commodity fetishism and a paradoxically fetishistic preoccupation with the traces and tokens of inner life. The novel **suggests** that the bourgeois subject, though it **comes** into being through its relations with things, **is defined** by the nostalgic notion that its true interiority **has been lost** under the pressure of things.

Text Sample 1.4a

Biochemistry text (Yamaguchi et al. 2005)

Verbs marked in **bold underline**

Several kinases **phosphorylate** vimentin, the most common intermediate filament protein, in mitosis. Aurora-B and Rho-kinase **regulate** vimentin filament separation through the cleavage furrow-specific vimentin phosphorylation. Cdk1 also **phosphorylates** vimentin from prometaphase to metaphase, but its significance **has remained** unknown. Here we **demonstrated** a direct interaction between Plk1 and vimentin-Ser55 phosphorylated by Cdk1, an event that **led** to Plk1 activation and further vimentin phosphorylation. Plk1 **induced** the phosphorylation of vimentin-Ser82, which **was elevated** from metaphase and **maintained** until the end of mitosis. This elevation **followed** the Cdk1-induced vimentin-Ser55 phosphorylation, and **was impaired** by Plk1 depletion.

It turns out that this difference in the density of verbs reflects an even more important difference in the typical syntactic structures of the two texts. At the clause level, the biochemistry text employs very simple syntactic structures, with few embedded dependent clauses. Some sentences consist of only a single main clause; for example:

*kinases **phosphorylate** vimentin. . .*

*Aurora-B and Rho-kinase **regulate** vimentin filament separation. . .*

Other sentences consist of two coordinated main clauses:

*Cdk1 also **phosphorylates** vimentin [. . .], but its significance **has remained** unknown*

*This elevation **followed** [. . .] phosphorylation, and **was impaired** by Plk1 depletion*

In fact, there are only two dependent clauses in the entire text passage from the science article. These are both relative clause structures modifying a head noun:

an event that **led** to Plk1 activation

the phosphorylation of vimentin-Ser82, which **was elevated** from metaphase [...]

In contrast, the literary criticism text is characterized by a high density of dependent clauses. In fact, every sentence in this text employs multiple dependent clauses, often with multiple levels of embedding. These include many different types of dependent clauses:

Complement clauses:

liked to **think** [...]

think that her novel "**touche[d]** [...]

reveal how **entrenched** [...]

suggests that the bourgeois subject [...] **is defined** [...]

Adverbial clauses:

published in the wake of the Great Exhibition

Though Bronte **liked** [...]

even as it **hopes** [...]

though it **comes** [...]

Relative clauses and noun complement clauses:

a world newly **captivated** by the peculiar resonance of things

an inner life that **eludes** [...]

terms that **reveal** [...]

an attempt to **negotiate** [...]

the nostalgic notion that its true interiority **has been lost** [...]

These two texts, which are typical of their respective disciplines, illustrate the fallacy inherent in the stereotype that all academic writing is the same. Rather than being homogeneous, consideration of actual text examples shows that there are systematic grammatical differences in the research

writing from different academic disciplines. When compared to conversation or popular written registers (like fiction or newspaper reportage), it makes sense to treat academic writing as a general register with distinctive grammatical characteristics. At the same time, though, there are systematic patterns of linguistic variation that distinguish the different types of academic writing. These differences are related to two other stereotypes about academic writing: the beliefs that academic prose employs complex and elaborated grammar, and that it is maximally explicit in meaning. We turn to those stereotypes in the following section.

1.2.2 Academic writing: Complex grammar and explicit meanings?

Grammatical features like the ones described in the last section contribute to the general perception that academic writing is more complex, structurally elaborated, and explicit in meaning than most other spoken and written registers. In fact, two of the most strongly held beliefs about academic writing are the stereotypes that it is extremely complex and elaborated in terms of its characteristic grammatical structures, and that it is maximally explicit in its presentation of information.

Hughes (1996: 33–34) notes these characteristics, writing that spoken grammar employs ‘simple and short clauses, with little elaborate embedding’, in contrast to written discourse, which employs ‘longer and more complex clauses with embedded phrases and clauses’, ‘explicit and varied marking of clause relations’, ‘explicit presentation of ideas’, and ‘explicit indication of text organization’.

The stereotypes that academic writing is elaborated and explicit are pervasive among writing teachers and researchers. For example, a simple database search of ERIC identified 114 published research articles where ‘writing’ and ‘elaborate/elaborated/elaboration’ appeared together, and 367 articles where ‘writing’ and ‘explicit’ occurred together. Although there is no consensus on the specific meanings of the terms ‘elaboration’ and ‘explicitness’ in these articles, there is clearly general consensus that these qualities can be attributed to academic writing as a whole, including both student writing and professional research writing. These views are often mentioned in passing, treated as background information rather than as a claim that requires evidence. Thus, consider the following quotes presented in Table 1.2 [with **emphasis added**].

These characterizations vary in the specific target of their descriptions (e.g., professional research writing or student academic writing). Further,

Table 1.2 *Characterizations of academic writing as elaborated and/or explicit*

Professional academic research writing is elaborated and/or explicit[...]	'[...] in academic writing [...] elaborated structures are generally preferred as they facilitate the readers' understanding of the text.' (Hyland & Tse 2005: 127)
...including across disciplines	'I have sought to make two broad claims in this paper: that elaboration is a complex and important rhetorical function in academic writing, and that <u>its use varies according to discipline</u>.' (Hyland 2007: 284) 'Here then we see the emphasis of the soft knowledge fields on the discursive exploration of possibilities and limiting conditions, identifying and elaborating relationships in argument.' (Hyland 2008: 11) 'This reflects the more discursive and evaluative patterns of argument in the soft knowledge fields, where persuasion is more explicitly interpretative and less empiricist [...]. The presentation of research is therefore altogether more discursively elaborate, [...]' (Hyland 2008: 16)
...increasingly so over time	'Moreover, the [historical] change to a more elaborate presentation of the data-analysis procedures may serve to strengthen the dependability (e.g. consistent, faithful, stable, unbiased.), accuracy (e.g. clearer, precise) and aptness (e.g., well-suited, appropriate, suited) of the findings to be reported subsequently in the Results section [...]' (Li & Ge 2009: 98)
Student academic writing that is high-quality and/or advanced is elaborated and/or explicit	'Students [writing chemistry lab reports] engage in elaborated discourse with a high degree of specificity [...]. Once they have focused on salient data and evidence, elaborated forms of discourse arrange information into more complex and explicit representations reflective of canonical scientific ideas.' (Wright 2008: 292) Myhill (1999) identifies elaboration and use of subordination as features which tend to characterise high quality Grade A writing [...]' (Keen 2004: 95) 'The redrafting process facilitates "progressively more extended clause planning and greater elaboration."' (Keen 2004: 96)

some of the descriptions are more nuanced and specific, identifying particular ways in which academic writing is elaborated and/or explicit. However, these descriptions all reflect the dominant and pervasive assumption that academic writing can be described as both elaborated and explicit.

For the researcher or teacher of academic writing, it would be nearly impossible to read current writing research and not form the general belief that academic writing is ‘complex’, ‘elaborated’, and ‘explicit’.

In descriptive linguistics, grammatical complexity and structural elaboration are strongly associated with the use of dependent clauses. By definition, a ‘simple’ clause has only a subject, verb, and object or complement. A ‘simple’ noun phrase has a determiner and head noun. Additions or modifications to these patterns result in ‘complex’ and ‘elaborated’ grammar. In particular, linguists have traditionally singled out dependent clauses as the most important type of grammatical complexity and structural elaboration (often contrasted with simple clauses or coordinated clauses; see, e.g., Huddleston 1984: 378; Willis 2003: 192; Purpura 2004: 91; Carter & McCarthy 2006: 489).

Building on this view, researchers in applied linguistics and second language writing development have operationalized the measurement of grammatical complexity in terms of the *T-unit*: a main clause and all associated dependent clauses. Wolfe-Quintero, Inagaki, and Kim (1998: 118–119) recommend variables like the number of dependent clauses per T-unit as the ‘best [...] complexity measures so far’, and numerous researchers in applied linguistics have applied similar measures in their research (see, e.g., Ellis & Yuan 2004; Brown et al. 2005; Larsen-Freeman 2006; Nelson & Van Meter 2007; Li 2000; Brown et al. 2005; Norrby & Håkansson 2007).

The preoccupation with dependent clauses as the best reflection of grammatical complexity leads to some bizarre conclusions when we consider actual texts from different registers. For example, conversational interactions regularly employ many embedded dependent clauses. Thus, consider the following two typical utterances from a casual dinner conversation:

Jill mentioned twice today
[that I need
[to find something
[to put my trophy on]]]

It was just one of those things
[where I think
[Paul’s gotten to the point
[where he won’t just go on and accept
[what she says]]]]]

It turns out that finite dependent clauses are actually more common in conversation than in academic writing (see discussion in Chapter 3). Thus, according to the stereotypical view of grammatical complexity, we

would be forced to conclude that conversation is more complex than academic writing – a conclusion completely at odds with the perceptions of any university student!

Of course, multiple embedded clauses do result in discourse that is more complex than alternative styles that rely on simple clauses, all other things being equal. And some written academic texts illustrate this type of grammatical complexity. For example, the literary criticism text discussed in the last section provides a good example of academic writing that employs extensive use of dependent clauses. In fact, all sentences in that passage incorporate multiple dependent clauses, often with multiple levels of embedding. Thus, consider the structure of the last sentence from [Text Sample 1.3](#), with three embedded dependent clauses (a verb complement clause, an adverbial clause, and a noun complement clause):

*The novel **suggests***

[that the bourgeois subject,

*[though it **comes** into being through its relations with things],*

***is defined** by the nostalgic notion*

*[that its true interiority **has been lost** under the pressure of things]]*

However, the stereotype that there is only one type of grammatical complexity – associated with clausal embedding – fails to capture the differences between conversational discourse and many sub-registers of informational written discourse. And the stereotype that all written academic texts are grammatically complex because they employ frequent dependent clauses is far from correct. For example, [Text Sample 1.4a](#) illustrated a passage from science research writing with surprisingly few occurrences of dependent clause structures. If we applied the traditional measures of grammatical complexity that focus on clausal embedding, we would conclude that this science text was not complex, because it employs little structural elaboration in the form of embedded dependent clauses.

So why do we have the perception that science writing is complex? Consider the first two sentences from [Text Sample 1.4](#). Both sentences have a simple clause structure with only a single main verb and no embedded dependent clauses:

1. Several kinases **phosphorylate** vimentin, the most common intermediate filament protein, in mitosis.
2. Aurora-B and Rho-kinase **regulate** vimentin filament separation through cleavage furrow-specific vimentin phosphorylation.

However, even though there are no dependent clauses, these sentences convey much embedded information. That information is expressed in embedded phrases rather than dependent clauses. However, the same information could be conveyed more explicitly through the use of multiple clauses. Thus compare 1b and 2b to the original sentences 1 and 2:

- 1b. *Vimentin is the most common intermediate filament protein; these proteins are structured to form intermediate filaments; kinases phosphorylate vimentin; this process occurs during the process of mitosis.*
- 2b. *Vimentin filament can be separated; Aurora-B and Rho-kinase regulate that separation; the separation occurs when something phosphorylates vimentin specifically where a furrow begins the process of creating cleavage*

The important point for our purposes here is to challenge the stereotype that all academic writing is grammatically complex because it employs frequent dependent clauses. Rather, [Text Sample 1.4](#) above illustrates a grammatical discourse style where information is conveyed through phrasal devices rather than through the use of dependent clauses. These phrasal features include nouns as a pre-modifier of another noun (e.g., *filament protein*, *vimentin filament separation*, *cleavage furrow-specific vimentin phosphorylation*), appositive noun phrases (e.g., *vimentin, the most common intermediate filament protein*), and prepositional phrases (e.g., *in mitosis*, *through [...] phosphorylation*).

These devices constitute a second major type of grammatical complexity, *not* associated with structural elaboration. In fact, they have exactly the opposite function: maximally compressing structure rather than elaborating structure. One of our primary goals in the following chapters is to document this type of grammatical complexity, describing the use of structurally compressed grammatical devices in academic prose (especially in science research writing). In the process, we directly challenge the stereotype that the complexities of academic writing involve extensive elaboration with frequent dependent clauses.

A related stereotype is the belief that academic writing is maximally explicit in the expression of meaning. These two stereotypes go together because it is commonly assumed that academic writing must be structurally elaborated in order to explicitly convey precise information. However, it turns out that the complexity devices actually preferred in science research writing (i.e., compressed phrasal devices) result in a major *reduction* in explicitness.

Clausal forms of expression are considerably more explicit than phrasal features, because they grammatically specify the meaning relationships among elements. For example, there are five meaning elements in the

noun phrase *cleavage furrow-specific vimentin phosphorylation*, but there are absolutely no overt grammatical signals to tell us the meaning relationships among those elements. Because we are not specialists in genetic biology, we are not certain about the intended meaning relationships among these five elements; our best guess is captured by the following clausal representation:

something ***phosphorylates vimentin specifically*** where a ***furrow***
begins the process of creating ***cleavage***

Of course, to be certain about the intended meaning relationships here, we would need advanced expert knowledge in biochemistry. Our point here, though, is to illustrate the compressed grammatical style of science research writing. There is an incredible amount of information compressed into the noun phrase *cleavage furrow-specific vimentin phosphorylation*, using only phrasal grammatical devices. And there are no explicit indications of the intended meaning relationships among the elements in that noun phrase. Thus, rather than being maximally explicit in the expression of meaning, it might be more accurate to portray certain types of academic prose (especially science research writing) as being maximally *implicit* in the expression of meaning. In [Chapter 6](#), we return to the lack of explicitness resulting from the use of phrasal grammatical devices, challenging yet another of the widely-held stereotypes about academic research writing.

1.2.3 Academic writing: Resistant to change or dynamic?

Another stereotype about academic writing is that it is conservative and therefore resistant to historical change, in contrast to spoken discourse, which is constantly in flux. In part, this stereotype is tied to beliefs about grammatical correctness as embodied in standard English. Academic prose is singled out as the register most likely to conform to the rules of standard English, in part because those rules are overtly taught in educational contexts. The norms of grammatical correctness are in turn assumed to be fixed and resistant to change, especially as they are codified in a written record:

Writing, however, is an artificial, conscious activity, and thus it is easy to resist language change in writing. We are taught to do just this, and therefore most written language is an artificial representation, omitting the signs of change which the real language, the spoken one, is full of. (McWhorter 2001: 17)

As a result, there is a widespread perception that written academic registers resist historical change and thus employ essentially the same grammatical style today as academic texts from a few centuries ago.

Hundt and Mair (1999) focus on academic writing as being especially resistant to change, in contrast to popular written registers that are ‘agile’ and receptive to change:

journalistic prose and academic writing seem to be two poles on a scale of written genres differing in the degree to which they are a) open to innovations or b) are prone to retain conservative forms. What we are proposing is a cline of openness to innovation ranging from “agile” to “uptight” genres’, with academic writing fitting on the ‘uptight’ end of the continuum [...] Because academic prose is less affected by economic pressures to win bigger audiences, it can afford to be more “old-fashioned”. (Hundt & Mair 1999: 236)

In contrast, spoken discourse is constantly in flux, and thus it has been the focus for research on historical change. Language change in the spoken mode is assumed to eventually bring about change in the written mode:

Indeed, writing slows language change down somewhat even on the spoken level, as writing reinforces our sense of “language” as a disembodyed blueprint to be followed or flouted No matter what the authority of the written form, or how tenaciously it holds on to the past, or how absurd the gulf between the written and the spoken form becomes, the spoken form always, always keeps on changing – and ultimately drags the written form reluctantly with it. (McWhorter 2001: 17)

This latter process can be documented in modern written academic texts that have begun to use innovative grammatical features adopted from speech. For example, the following sentences from modern academic research articles might appear to be unremarkable, but they each contain a grammatical feature that would have been virtually unattested in academic writing from earlier centuries:

1. *Orwell, whether he knew it or not, was actually advocating critical support for Churchill's war government. In truth, Macdonald was being a bit disingenuous.*
2. *The exact meaning of “fixation” on the molecular level will have to be settled, however, by the study of well defined molecular species.*
3. *Then the nucleic acid of the particle becomes released and reaches the cellular site where its replication is going to take place.*

These sentences illustrate grammatical innovations that emerged in speech in recent centuries: the more widespread use of progressive verbs (especially progressive stative verbs, like *being*) in #1; the semi-modal *have to* in #2; and the semi-modal *be going to* in #3. Such features reflect the general historical trend towards more ‘colloquial’ styles over the last

century, where grammatical innovations from spoken discourse are becoming increasingly prevalent in written discourse (see, e.g., Hundt & Mair 1999; Mair 2006; Leech et al. 2009). This historical drift has especially affected popular written registers like fiction and newspaper writing (see the corpus analyses in Chapters 3 and 4). Academic research writing has for the most part resisted the adoption of these colloquial features. While features like stative progressive verbs and semi-modals can be traced back to the eighteenth and early nineteenth centuries in popular written registers, they have only recently begun to appear in academic writing (and they are still extremely rare in that register). Thus, although it is possible to find instances of innovative colloquial features in academic writing, the evidence in this regard overwhelmingly supports the stereotype that academic writing is conservative and resistant to historical change.

In other respects, though, academic writing has been at the forefront of historical change. These grammatical innovations do not involve the use of stigmatized colloquial or non-standard forms, and thus they have generally not been noticed by usage commentators. And historical linguists have assumed that linguistic change originates in speech (and then spreads to popular written registers), and thus they have generally disregarded the possibility of historical change in the use of other grammatical devices that originates in academic writing. However, empirical corpus-based research shows that the grammatical discourse style of academic writing – especially science research writing – has changed radically over the past 300 years. One of our primary goals in the present book is to document those patterns of change.

To illustrate these historical developments, consider the following nineteenth century academic texts, comparing their grammatical characteristics to the modern humanities and science texts discussed in Sections 1.2.1 and 1.2.2. Text Sample 1.5 is taken from an 1852 book on the early history of Otsego County in New York, while Text Sample 1.6 is from a science article published in the *Philosophical Transactions of the Royal Society of London* in 1800.

Text Sample 1.5 Nineteenth century history

Discarding all claims to literary merit, I hope to contribute something that shall instruct and amuse the rising generation; the tendency of which, shall be to make them better citizens, and fit them to discharge the various duties of life, with credit to themselves and benefit to our common country. Individual incidents unimportant as they appear, go to make up when combined, what we call history; and though generally over-looked, except as connected with the great men of the earth,

yet as little rills and streamlets feed the mighty rivers, that sweep along with irresistible force; so individual events help to make up the great historical mass, that floats down the current of time.

Levi Beardsley. 1852.

'Reminiscences: Personal and other incidents; Early settlement of Otsego County; Notices and anecdotes of public men; ...'

Published by Charles Vinten.

Text Sample 1.6 Nineteenth century science

The discoveries which have been made with respect to light, as it proceeds immediately from the sun, are many and important; but the observations on that species of light which is spontaneously emitted from various bodies, are not only few in number, but in general very imperfect. The author is therefore desirous of drawing the future attention of the philosopher more particularly to this subject, and of communicating his own experiments and observations upon it, to this learned Society.

By the spontaneous emission of this light, the author wishes to distinguish it from all kinds of artificial phosphorus; which, as he apprehends, differ essentially, in some of their properties from that light of which he means to treat. And, by its adhesion to bodies with some degree of permanency, he distinguishes it from that transient sort of light which is observable in electricity, in meteors, and in other lucid emanations. The light which is the subject of this paper, he shall therefore beg leave to discriminate by the name of spontaneous light.

Nathaniel Hulme. 1800.

'Experiments and Observations on the Light which is spontaneously emitted, with some Degree of Permanency, from various Bodies'

Philosophical Transactions of the Royal Society of London, 90: 161–187.

These nineteenth century academic texts are similar grammatically to present-day humanities texts. For example, the nineteenth century history text illustrates a dense use of attributive adjectives; in fact, almost every noun phrase is modified by an attributive adjective in this passage:

Attributive adjectives:

literary merit

rising generation

better citizens

various duties

common country

individual incidents

great men

little rills and streamlets

mighty rivers

irresistible force

individual events

great historical mass

The nineteenth century science text is also similar to modern humanities writing in its use of attributive adjectives as well as frequent nominalizations:

Attributive adjectives:

various bodies

learned Society

artificial phosphorus

spontaneous light

Nominalizations:

discoveries

observations

adhesion

permanency

Attributive adjective + nominalization:

future attention

spontaneous emission

lucid emanations

At a more specific level, the preferred syntactic structures employed in the nineteenth century history article ([Text Sample 1.5](#)) are in many

ways similar to those used in the modern literary criticism article (Text Sample 1.3), especially with respect to the use of long sentences incorporating multiple dependent clauses; these include all major types of dependent clauses:

Complement clauses:

I hope to contribute something . . .
shall be to make them better citizens. . .
make up [...] what we call history. . .
help to make up the great historical mass . . .

Adverbial clauses:

Discarding all claims. . .
to discharge the various duties. . .
to make up. . .
when combined. . .
though generally over-looked
except as connected with the great men . . .
as little rills and streamlets feed the mighty rivers . . .

Relative clauses:

something that shall instruct and amuse. . .
the tendency of which, shall be. . .
the mighty rivers, that sweep along . . .
the great historical mass, that floats down the current . . .

The main difference between the nineteenth century history text and the modern humanities text is the extent to which they rely on dependent clauses: the nineteenth century humanities text is marked by this characteristic to an even greater extent than modern humanities prose. The result is often incredibly long sentences that are extremely complex in terms of their clausal embedding. Thus, consider the clausal structure of the last sentence from Text Sample 1.5:

Individual incidents, go
 [to make up
 [when combined]
 [what we call history]]
 and
 [though generally over-looked
 [except as connected with the great men of the earth]]
 [yet as little rills and streamlets feed the mighty rivers,
 [that sweep along with irresistible force]]
 so individual events help
 [to make up the great historical mass,
 [that floats down the current of time]]

This comparison of nineteenth century and twentieth century humanities texts shows that there has been historical change in the *extent* to which elaborated structures (with multiple dependent clauses) are used, but there is little evidence of change in the *types* of grammatical structures that are employed. In contrast, a comparison of the nineteenth century science article in [Text Sample 1.6](#) and the modern science article in [Text Sample 1.4](#) illustrates historical change at a much more fundamental level. The nineteenth century science text is similar to humanities texts (both nineteenth and twentieth centuries) in utilizing extensive clausal embedding. [Text Sample 1.6](#) particularly relies on relative clauses, but adverbial clauses are also used:

Relative clauses:

The discoveries which have been made. . .
that species of light which is spontaneously emitted. . .
all kinds of artificial phosphorus; which, as he apprehends, differ essentially. . .
that light of which he means to treat
that transient sort of light which is observable in electricity. . .
the light which is the subject of this paper. . .

Adverbial clauses:

as it proceeds immediately from the sun
as he apprehends

Table 1.3 *Summary of major grammatical differences between four text excerpts*

	Nineteenth Century	Present Day
Humanities	<i>Text Sample 1.5: history</i> + attributive adjectives ++ clausal embedding	<i>Text Sample 1.3: literary criticism</i> + attributive adjectives + clausal embedding
Science	<i>Text Sample 1.6: physics</i> + nominalizations + clausal embedding (relative clauses) - nouns as noun pre-modifiers - appositive noun phrases	<i>Text Sample 1.4: biochemistry</i> + nominalizations - clausal embedding ++ nouns as noun pre-modifiers + appositive noun phrases

The density of these dependent clause structures is not nearly as great as in the nineteenth century history text, but the general tendency is similar: to structure discourse through the use of embedded dependent clauses. This tendency is especially noteworthy when we contrast the grammatical style of this nineteenth century science text with the preferred grammatical structures in [Text Sample 1.4](#) (the modern science research article). On the one hand, the nineteenth century science text relies on dependent clause structures to a much greater extent than the modern science text. But the more interesting difference concerns the grammatical features that are not found in the nineteenth century text: nouns as pre-modifiers of a head noun, or appositive noun phrases post-modifying a head noun. These features represent important grammatical innovations that are found in modern science research writing, but are generally much less common in other written registers (including modern humanities academic writing).

[Table 1.3](#) summarizes the trends that we have identified so far, comparing two parameters of variation: discipline (humanities vs. science) and historical period (nineteenth century vs. modern day). While some linguistic features are consistently used in academic writing across disciplines or across historical periods, others show marked differences along these same parameters. In particular, the table highlights the peculiar nature of modern science research writing, strikingly different from present-day humanities writing as well as nineteenth century science writing. This register is the locus of the types of grammatical change documented in the present book. As summarized in [Table 1.3](#), grammatical features like pre-modifying nouns and appositive noun phrases were rare in all registers in the nineteenth century. In contrast, these phrasal devices are so prevalent in present-day academic science writing that they usually escape notice.

The present book describes the distinctive grammatical characteristics of modern science research writing, and documents the dramatic patterns of historical change towards that grammatical style. Those changes include a strong decrease in the use of dependent clauses, accompanied by a strong increase in the use of phrasal modifiers. As a result, modern science writing regularly employs phrasal complexity features that are minimally used in earlier historical periods (e.g., pre-modifying nouns and appositive noun phrases). These historical developments are of particular interest because they contradict generally accepted claims about historical change in English. In particular, we show that

1. grammatical innovation and change occurs in writing as well as in spoken interaction; and
2. academic written registers are *not* conservative and resistant to change.

Thus, the analyses reported in the following chapters show how certain kinds of grammatical features have expanded historically in their frequency of use, their grammatical productivity, and their meanings and functions – and how all of these changes have occurred primarily in academic writing (especially science research writing), reflecting the special communicative characteristics of that register.

1.3 Grammatical change in English: Below the radar?

Historical change in English is an important topic treated in almost all introductory textbooks on linguistics and sociolinguistics. For the most part, though, these treatments focus on changes in pronunciation/phonology or lexical change. For example, several books document phonological changes in English like the loss of initial [h] in ‘which’, ‘whether’, and ‘whine’, or lexical changes like *chesterfield* > *couch* and *serviette* > *napkin* (see, e.g., Holmes 1992; Chambers 2002). In many cases, these changes result in present-day patterns of regional or social dialect variation. Examples of this type include the loss or retention of post-vocalic [r], the merger of [t, d] with [θ, ð], and the vowel shifts found in American northern cities (see, e.g., Wolfram and Schilling-Estes 2005).

Grammatical change is generally disregarded in introductory linguistics textbooks. The few books that deal with this topic at all tend to focus on morphological change, such as the regularization of verb paradigms for irregular verbs (e.g., *dreamt* → *dreamed*), or vice versa (*sneaked* → *snuck*). Some books also discuss patterns of morphological change (or resistance to change) that have resulted in present-day dialect differences (see, e.g.,

Wolfram and Schilling-Estes 2005: 85–92; Finegan 2008). Such changes include the loss of third person singular verb inflection (*he runs* → *he run*) and the development of second person plural pronouns in many regional dialects of English (e.g., *y'all*, *youse*, *joins*).

In this regard, the coverage in linguistics textbooks is typical of linguistics research generally, which places ‘much more emphasis on phonological and lexical phenomena than on grammatical ones’ (Leech et al. 2009: 7). One explanation for this fact is that we just fail to notice grammatical change: that ‘we are considerably less perceptive of it than of other kinds of linguistic change’ (Strang 1970: 60, cited in Leech et al. 2009: 7). Leech et al. go on to discuss the difficulty in noticing grammatical changes. One obvious factor is the nature of the phenomena. When the phonetics of a word begins to change, we notice that some people pronounce the word differently from others. When a new word comes into the language, we notice that some people are using the word and others not. From a synchronic perspective, such changes usually manifest themselves as social patterns of dialect variation; consequently, we readily notice some groups of speakers using pronunciations and words that are not used by other groups of speakers. Such patterns of variation can evolve rapidly, so that there are often dramatic phonological and lexical differences from one generation of speakers to the next (see, e.g., Labov 1994, 2001, 2010).

In contrast, grammatical change is usually more abstract, involving the use of different phrase types and different clause types, and changes in the ways in which those structures are configured relative to one another. Probably an even more important consideration is that grammatical change proceeds much more slowly than lexical/phonological change, and it is less clearly associated with dialect variation.

The grammatical changes that have been most often noticed are typically lexico-grammatical: changes in the grammatical function of specific words. Prescriptive grammarians often target specific lexico-grammatical features associated with the general ‘decay’ of written English. Most of these pundits focus on present-day use and perceived misuse, but the targeted features are often specific words undergoing grammatical change. Examples include variation in the past participle form for irregular verbs (e.g., *dove* vs. *dived*, *learnt* vs. *learned*, *lit* vs. *lighted*), preposition choice (e.g., *different from* vs. *different than*), *less* versus *fewer* with count nouns, *like* functioning as a subordinator or a preposition, the conversion of adjectives to adverbs (e.g., *real good*; *go fast/slow*), and the choice between *that*, *who*, and *whom* as relative pronouns with an animate head noun.

Similarly, sociolinguistic studies of grammatical change usually focus on developments that affect specific words. For example, numerous studies

have documented the recent grammatical change associated with features like the new quotative verbs (e.g., *go*, *be like*, *be all*), the discourse marker *like*, and *so* used as an intensifier modifying an adjective (see the survey of research in Tagliamonte 2006). Changes of this type affect only a few specific words, they occur in a relatively short time span, and they are usually associated with strong patterns of social dialect variation; as a result, such changes are often noticed and described by professional linguists as well as the wider public.

In contrast, other grammatical changes are general structural shifts, influencing an entire construction type or grammatical system. For example, over the last few centuries there have been general changes in the use of modal verbs (e.g., Leech 2003, 2011; Millar 2009), verb phrases with marked aspect (e.g., Smith & Rayson 2007; Nesselhauf 2007; Mair and Hundt 1995; Smith 2002; Hundt 2004; Mair 2006), and complementation patterns (e.g., Mair 1995, 2002; Rudanko 2010). Such changes have a lexical aspect, in that they gradually work their way through the lexicon. However, these are not innovations restricted to a few individual words. Further, in contrast to the restricted lexico-grammatical changes discussed in the last paragraph, these general structural changes occur gradually over the course of several centuries, and they typically do not serve to distinguish among social dialects during their period of transition. As a result, such historical developments have attracted much less attention than short-term, socially-stigmatized grammatical changes in the use of particular words.

For example, progressive aspect has increased strongly in use over the past three centuries, developing new meanings and functions which have gradually occurred with an ever increasing set of lexical verbs (see Hundt 2004; Mair 2006; Leech et al. 2009). If we compare the end points of this historical change, the differences are as dramatic as the recent development of new quotative verbs like *go* and *BE all*. However, historical change in the use of progressive aspect has been much more gradual, occurring over a period of centuries rather than a period of a few years. And there is no particular verb that we can point to as the target of this change; rather, the use of progressive aspect has been generally increasing for an ever expanding set of verbs.

These two characteristics – long-term change (vs. rapid change) and change influencing a grammatical system (vs. change that targets a few specific words) – are typical of many grammatical developments in English. Such patterns of change are not salient to speakers of a language at any given point in time, and they are not obviously associated with synchronic patterns of dialect variation. As a result, linguistic change of

this type has generally been disregarded in comparison to lexical and phonological changes. However, several recent books on the history of the English language include descriptions of such grammatical change. In the following section, we survey some of the major features discussed in those treatments.

1.4 Two types of grammatical change

A long-term perspective on the history of English makes it clear that there have been numerous important grammatical changes in English over the course of the last 1,000 years. These include the change to a fixed SVO word order, the loss of most inflectional morphology (especially case suffixes), the increase in the range of function words (including prepositions, auxiliary verbs, and infinitive marker *to*), and more recently the introduction of the dummy auxiliary DO (see Rissanen 1999; van Gelderen 2006).

However, grammatical changes over the last 300 years – the period of Modern English – have been much less dramatic. There have been numerous interesting developments, but they differ in kind from the earlier dramatic changes. Recent historical developments include many of the features that have been discussed by prescriptivists, such as changing preferences for irregular versus regular verb forms, the choice of *fewer* versus *less*, and the decline in use for the relative pronoun *whom*. However, there have also been changing patterns of use for numerous other grammatical features, documented in survey articles by Barber (1964: 130–144), van Gelderen (2006: 171–174, 214–220), and Fennell (2001: 144–145, 173–174), as well as in the more extended treatments by Denison (1998), Mair (2006), and Leech et al. (2009). Some of the more important changes are:

- development of auxiliary-like uses for *want to* → *wanna*
- increase in the use of the semi-modals *be going to* (*gonna*) and *have got to* (*gotta*)
- decrease in use of the core modal verbs
- increasing use of analytical rather than synthetic comparison
- increasing use of the *s*-genitive, especially with non-human nouns
- decreasing use of the *of*-genitive
- increasing use of the ‘mandative’ subjunctive
- decrease in the use of *shall* as a future marker
- increasing use of progressive constructions

- development of the progressive-passive
- development of the *get* passive
- overall decreasing use of the *be* passive
- increasing tendency to omit the relative pronoun *that*
- increase in the number and type of multi-word verbs
- changing complementation patterns with non-finite clauses

These linguistic developments illustrate two major types of grammatical change:

- a) the emergence and development of new grammatical features (including new grammatical uses of an individual word); and
- b) shifts in the typical use and functions of core grammatical features

The first type of grammatical change has been studied mostly under the rubric of ‘grammaticalization’, which focuses on the way in which content words evolve over time to be used as grammatical function words. Examples from the list mentioned earlier include the use of *have to* and *got to* as semi-modals, *wanna* with modal auxiliary functions, and *get* as an auxiliary verb in passive constructions. Other examples include the use of *well* as a discourse marker, *pretty* as a hedge or intensifier, and sequences like *in spite of*, *with regard to*, and *because of* used as complex prepositions (see, e.g., Krug 2000; Tagliamonte 2004; Hopper & Traugott 2003; Lindquist and Mair 2004; Nevalainen 2004; Hoffmann 2005).

The new quotative verbs illustrate a related kind of grammatical change. Some of these forms – like *go* – have evolved from a lexical verb to be used in the more grammaticalized function of marking reported speech and thought. The quotative *be all* also seems to have evolved from a lexical (adjectival) use to this more grammaticalized function. However, it could be argued that *be like* evolved in the opposite direction, from a function word (a preposition or subordinator) to a more lexicalized grammatical function as a quotative marker. Regardless, these are all examples of grammatical innovation that result in particular words changing to be used with new grammatical functions.

In contrast, the second type of grammatical change is completely different in kind: it focuses on the extent to which a grammatical feature is used rather than the development of new grammatical classes. This type of grammatical change affects the overall frequency of a feature and the co-occurrence of the grammatical feature with an ever increasing set of words; these changes are also often associated with an increase in the meanings and functions expressed by the grammatical feature. Examples

of this type from the above list include the increasing use of the *s*-genitive, progressive verbs, and multi-word verbs. At the same time, other features have declined in use and functionality over the past few centuries (e.g., modals and passive voice verbs).

It turns out that most grammatical change in English over the past 300 years has been quantitative rather than categorical. Even grammatical innovations, like the development of semi-modals and the *get*-passive, have gradually continued to increase in frequency and functionality over this period. Thus, a quantitative perspective on grammatical change has become increasingly important in recent research:

...changes in the realm of syntax are often a function of quantity, rather than quality; that is, certain structures have expanded in number and frequency of occurrence during the PDE period. (Fennell 2001: 173)

Since relatively few categorical losses or innovations have occurred in the last two centuries, syntactic change has more often been statistical in nature, with a given construction occurring throughout the period and either becoming more or less common generally or in particular registers. The overall, rather elusive effect can seem more a matter of stylistic than of syntactic change. . . (Denison 1998: 93)

Most of the diachronic changes that we document in this book are of this type, relating to the frequency of occurrence and the functionality of grammatical features, rather than the development of completely new grammatical categories. It turns out that grammatical change of this type has been especially prevalent in academic writing, unlike the lexicogrammatical innovations typical of speech.

1.5 The locus of historical change in English: Speech or writing?

It will be clear from the preceding sections that we are interested primarily in grammatical innovations that have occurred in the discourse of academic writing. For many historical linguists, this focus is a non-starter. That is, most previous research on grammatical change has focused on spoken discourse, and it has often overtly excluded the possibility of grammatical change in writing.

The view that language change occurs in speech – and not in writing – can be attributed to several sources. At the most basic level, this view reflects the widespread belief in linguistics that spoken language is ‘real’, while written language is a derivative artifact (see discussion in Biber 1988: 5–7). In the early American linguistic tradition, Sapir and Bloomfield both held this position:

writing is ‘visual speech symbolism’ (Sapir 1921: 19–20)

writing is ‘not language, but merely a way of recording language by visible marks’ (Bloomfield 1933: 21)

And numerous other linguists in more recent decades have reiterated this position:

‘speech is fundamental and writing . . . only a secondary derivative’ (Hall 1964: 8–9)

Written communication is ‘derivative of the face-to-face conversational norm’ (Fillmore 1981: 153)

‘spoken language is “true” language, while written language is an artifact’ (Aronoff 1985: 28)

‘Language, Saussure taught us, is first and foremost a spoken system – writing is a secondary coding, but speech is primary’ (Schneider 2002: 67 – citing Saussure 1916)

‘Writing, however, is an artificial, conscious activity, and thus it is easy to resist language change in writing. We are taught to do just this, and therefore most written language is an artificial representation, omitting the signs of change which the real language, the spoken one, is full of.’ (McWhorter 2001: 17)

Nearly all introductory textbooks on linguistics continue to reflect this bias in their reference to ‘speakers’ and ‘speech communities’, with little or no acknowledgement of ‘writers’ or ‘writing communities’.

Historical linguists who study recent change in English have similarly been preoccupied with the characteristics of speech, describing the ways in which English has become increasingly colloquial and informal in recent centuries. The linguistic features associated with these changes originated in speech, and they have gradually been adopted in writing. Thus, to the extent that historical researchers have considered change in writing, they have focused mostly on the increasing use of colloquial features borrowed from speech, such as first person pronouns, contractions, and semi-modals. This historical trend towards a more colloquial style has been documented by numerous discourse analysts and corpus linguists, and is referred to as the ‘drift’ of written registers towards more ‘oral’ styles (Biber and Finegan 1989a), ‘informalization’ (Fairclough 1992), and ‘colloquialization’ (Hundt and Mair 1999; Mair 2006; Leech et al. 2009). For example, Mair (2006) describes this process as follows:

For some time now, for example, and increasingly so since the 1960s and 1970s, an egalitarian and informal communicative culture has

been promoted in the public domain which has brought the norms of writing closer to the norms of spoken usage. In grammatical terms, this has favored the rapid disappearance of archaisms such as *upon* for *on* or the subjunctive in all but its mandative use, and led to a decrease in the popularity of typical markers of formal and written style such as the passive voice. On the other hand, it has facilitated the spread of informal grammatical options such as contractions, the *going to*-future, or certain types of progressives into domains in which they used to be rare. (Mair 2006: 88)

Earlier investigations of these patterns suggested that they represent a general historical trend in English. However, subsequent investigations show that there are important differences among written registers. For example, Biber and Finegan (1997/2001) show that written registers like fictional novels and personal letters have been strongly influenced by the drift towards more colloquial linguistic styles, but written academic registers (especially science research articles and medical research articles) have not participated in that trend. Hundt and Mair (1999) also note this difference, distinguishing between ‘agile’ written registers (e.g., newspaper prose) that are receptive to these changes, and ‘up-tight’ written registers (e.g., academic prose) that resist historical change towards the increased use of colloquial features.

A complementary approach to historical change has focused on the ways in which grammatical innovations emerge in language use, carried out under the rubrics of ‘grammaticalization’, ‘usage-based approaches’, ‘emergent grammar’, and the general study of ‘frequency effects’ in language use. These approaches share a focus on grammatical use in natural discourse, describing how new grammatical constructions and/or functions emerge from natural communicative situations.

Most usage-based studies of grammatical change focus on spoken interaction, using the term ‘speaker’ as a cover term for the addressor or producer of discourse, and often explicitly framing the discussion in terms of ‘utterances’ and conversational ‘interaction’.¹ Written discourse has generally received little attention, with most studies simply disregarding the possibility that grammatical innovations might develop in natural written communication.

For example, an exclusive focus on spoken interaction is front-and-center in the edited book on *Interaction and Grammar* by Ochs, Schegloff,

¹ A few usage-based studies acknowledge the possibility of historical change in writing as well as speech. For example, Traugott (2003: 125) mentions the possible role of the ‘writer’ in her definition of ‘subjectification’: the tendency of meanings ‘to become increasingly based in the SP[eaker]/W[riter]’s subjective belief state or attitude to what is being said and how it is being said’.

and Thompson (1996), which approaches grammar as part of the social practices associated with conversational interaction. The studies in this book analyze transcripts of conversations, showing how grammar emerges with special functions in spoken interactional contexts. A similar implicit emphasis on spoken interaction can be observed in so-called usage-based models of grammar (see, e.g., Langacker 1987; Kemmer & Barlow 2000). While these descriptions tend to be based on intuitive notions of spoken interaction, rather than direct analysis of conversational texts, they still implicitly frame the discussion relative to spoken discourse; for example, the ‘usage events’ that form the foundation of the usage-based model are ‘utterances’ produced by ‘speakers’ (Langacker 2000: 9).

Similarly, in describing the motivations and enabling factors of grammaticalization, Hopper and Traugott (2003: 71) note that previous research has focused on ‘the role of speakers and hearers negotiating meaning in communicative situations’. Bybee and Hopper (2001) are also typical in their focus on spoken interaction, noting that:

The notion of emergence . . . relativizes structure to speakers’ actual experience with language, and sees structure as an on-going response to the pressure of discourse. . .

The distribution and frequency of the units of language are governed by the content of people’s interactions. . .

Patterns of use . . . deal with patterns of occurrence of morphosyntactic structures in natural conversation. . . (Bybee & Hopper 2001: 3)

Croft (2000) is especially emphatic in arguing that language change occurs in utterances produced by speakers in conversational interaction; for example:

language use is essentially a joint act between speaker and addressee
 . . . Language is a fundamentally social interactional phenomenon.
 So is language change. (Croft 2000: 87)

Given this background, it is not surprising that many studies of grammaticalization have focused on grammatical features that are common in conversation but rarely used in writing, such as the English semi-modals (e.g., *have to*, *got to*), discourse markers (e.g., *well*, *like*), intensifiers (e.g., *so*), and quotative verbs (e.g., *go*, *be all*).

As noted above, most scholars simply disregard the possibility that grammar might also emerge in written use. One of the few exceptions is Croft (2000), who directly considers the possibility but then strongly argues against it:

One might speculate that the advent of the written medium led to directed evolution in the development of these construction types [e.g., nominalizations, participles, attributive adjectives] . . . Typological research indicates, however, that all of these construction types are present in most if not all unwritten languages. (Croft 2000: 83)

It is thus possible that there is directed change in the advent of the written register. However, expansion into the new linguistic niche results in the evolution at most of new degrees of syntactic complexity – multiple iterations and embeddings of structures – rather than in developing completely new grammatical structures. (Croft 2000: 83–84)

In this passage, Croft is specifically discussing grammatical innovation from a typological perspective: can a new type of grammatical construction emerge in a written language in comparison to the types of constructions that already exist in the world's spoken languages? However, his basic position is in line with the generally accepted view of most historical linguists: that grammatical innovations occur in speech, not writing; and that any grammatical developments in writing are derivatives from speech.

Of course, historical linguists do not have access to speech samples from earlier historical periods, and thus they are forced to rely on the analysis of written texts. It might be supposed that this methodological requirement would have forced linguists to consider the possibility of grammatical change originating in writing. But that has not occurred. In fact, the methodological focus of historical research has been on ways to filter out the assumed 'obscuring' effect of writing, in the attempt to uncover the 'true' patterns of language use that occurred in speech from earlier periods.

Schneider (2002) has written one of the most thorough treatments of this methodological issue, under the promising title 'Investigating variation and change in written documents'. This title might lead the reader to expect discussion of the importance of grammatical change in writing. However, the very first sentence of the article immediately quashes that possibility, asserting that speech is primary and writing merely 'a secondary coding' (Schneider 2002: 67). Schneider goes on to discuss the implications of this view for historical research:

Normally, as variationist linguists we are not directly interested in the written record as such [...] – its function is predominantly to serve as a clue, a pathway to the variation and change of the language system in itself. More directly, most written records of interest in this context represent a speech act [...]. In such cases, the written record functions as a filter: [...] the rendering of the speech event is only indirect and imperfect. (Schneider 2002: 67)

A written record of a speech event stands like a filter between the words as spoken and the analyst. As the linguist is interested in the speech event itself [. . .], a primary task will be to “remove the filter” as far as possible. (Schneider 2002: 68)

In sum, Schneider argues that historical linguists have no choice but to analyze written documents. However, they should have no expectation of uncovering grammatical changes that emerge in writing. Rather, the focus is exclusively on speech (representing the real ‘language system’), and the methodological challenge is therefore to uncover the patterns of use that existed in speech, based (unfortunately) on the analysis of written texts.

Culpeper and Kytö (2010) directly address this methodological challenge by carrying out detailed linguistic analyses of spoken dialogue as represented in speech-based registers from Early Modern English (e.g., legal trial proceedings, witness depositions, drama). Other researchers (e.g., Nevalainen and Raumolin-Brunberg 1996; Denison 1998) have collected texts from informal/colloquial registers, like personal letters or diaries, in the hope that these will approximate the underlying linguistic system found in speech from earlier historical periods. These studies have been highly important for the wealth of detailed information that they provide about the typical linguistic patterns of use in earlier periods. However, they also reflect the dominant view that the primary inquiry of interest in historical linguistics concerns the patterns of change that occurred in speech. Written records allow us to infer what those changes have been, but they are assumed to be of no interest in themselves. And in particular, the language of informational written prose is considered to be far removed from the kinds of discourse that historical linguists should be analyzing to uncover genuine historical change.²

In the present book, we take a fundamentally different position, arguing that the communicative demands of any register – whether spoken or written – have the potential to facilitate the emergence of grammatical uses (and possibly new constructions) associated with those communicative needs. The first author has been advocating this position for the last 25 years; for example:

² Our characterization of the state-of-the-art here disregards the long tradition of research by historical corpus linguists such as Rissanen, Nevalainen, and Crespo, who have carried out numerous historical studies of corpora consisting of written documents (specifically written academic documents in the case of many of the corpora constructed at the University of Helsinki and the University of A Coruña).

...the two modes of communication [speech and writing] have quite different strengths and weaknesses, and they therefore tend to be used in complementary situations. From this perspective, neither can be said to be primary: they are simply different. The linguistic characteristics of each mode deserve careful attention, and the relationship between the two modes must be investigated empirically rather than assumed on an a priori basis. (Biber 1988: 9)

Biber and Conrad (2009: 6) further develop this argument, claiming that 'linguistic features are always functional when considered from a register perspective'; that 'linguistic features tend to occur in a register because they are particularly well suited to the purposes and situational context of the register'. These functionally-motivated patterns of linguistic variation characterize present-day spoken registers (Chapter 4) and written registers (Chapter 5), as well as historical change in writing (Chapter 6).

Fox (2007: 299) has also advocated investigation of this possibility, stating that: 'The relationship of written language to language-in-interaction is complex and worthy of independent study'. Thus, Fox argues that written discourse is produced under completely different circumstances from spoken discourse, and that these characteristics might have direct grammatical consequences:

Writers and readers typically have no time constraints placed on their production and comprehension, a fact which presumably allows more complex syntactic structures to arise ... On the other hand, speakers and recipients in real-time conversation have immense time pressures on them ... [creating] a tendency in conversation towards shorter and syntactically simpler utterances. ... (Fox 2007: 314)

It is clear ... that writing alone, at leisure ... is a different grammatical enterprise than is designing an utterance, in real time ... the grammar of written language thus needs to be taken up as a separate investigation. ... (Fox 2007: 315)

The present book explores this possibility, based on the underlying assumption that the communicative circumstances of any register will provide the context for functionally-motivated grammatical change. In particular, we focus on academic research writing as a register that differs in almost every way from face-to-face conversation:

- written rather than spoken
- monologic rather than interactive and co-constructed
- requiring specialized, professional background knowledge, but no assumption of personal background knowledge
- slowly produced and carefully revised and edited

If grammatical change is influenced by the pressures of the communicative situation, there is every reason to expect that the grammatical devices and functions emerging in written academic writing will be strikingly different from those that have emerged from conversational interactions.

The analyses in the following chapters show convincingly that historical written texts should not be treated as a ‘filter’ that obscures true historical change. Rather, extensive grammatical change has occurred in writing. These patterns of change have affected different grammatical characteristics from the colloquial features described in most previous research. However, the grammatical innovations of written discourse cannot be discounted as mere ‘iterations and embeddings’. In particular, focusing on noun phrase constructions in English, the following chapters document extensions in the range of grammatical variants, the range of lexical associations, and the range of grammatical/semantic functions – in addition to dramatic quantitative changes in the frequency of use.

1.6 Overview of the present book

The present book challenges previous assumptions about grammatical complexity, historical linguistic change, and academic writing, documenting the ways in which academic writing has been linguistically innovative over the past two centuries. The primary focus of the book is the description of phrasal complexity features and the associated phrasal discourse style that is typical of present-day science research writing. From a synchronic perspective, the book documents how academic written registers are unlike other registers – spoken or written – in their reliance on phrasal complexity features rather than a clausal style of discourse. From a diachronic perspective, the book documents linguistic innovations that have emerged in writing (especially science research writing) but have generally not been adopted in speech.

We base our descriptions on large-scale corpus analysis that allows us to track long-term change in the grammatical structure and discourse style of academic writing, as compared to a range of other written and spoken registers. [Chapter 2](#) introduces the corpus-based approach to linguistic analysis, as well as the specific corpora and analytical methods used for our study. It begins with a discussion of recent trends and issues in diachronic corpus-based research. The chapter then introduces the corpora that serve as the basis for our analyses; these span 300 years and represent a range of registers (e.g., everyday conversation, newspaper writing, fiction, academic writing). The academic writing corpus is designed to represent several sub-registers, including humanities prose, social science articles, and natural science articles.

Chapter 2 also outlines the research methods for our analyses, including corpus annotation, interactive and automatic analysis techniques, and quantitative research designs. Finally, Chapter 2 introduces and exemplifies the grammatical features discussed in the book, including the full set of clausal and phrasal complexity features.

Chapters 3 through 6 present the major findings of our research. Chapters 3 and 4 present quantitative corpus findings, while Chapters 5 and 6 are more interpretive, presenting the results of detailed linguistic analyses. These chapters are sequenced as follows: Chapter 3 describes synchronic patterns of register variation; Chapter 4 presents the results of similar analyses carried out from a diachronic perspective; Chapter 5 presents detailed linguistic analyses to support the claim that these historical developments are not ‘merely’ stylistic changes in frequency; and finally, in Chapter 6, we turn to the question of explicitness, showing how phrasal complexity features are less explicit in meaning than alternative clausal features.

Chapter 3 provides a synchronic description of the grammatical discourse style of academic writing, focusing especially on features associated with grammatical complexity and elaboration, but also describing the use of common, core grammatical features. A major goal of this chapter is to show that traditional operational definitions of ‘complexity’ – in terms of clausal embedding – fail to capture the actual complexities of academic writing, which are phrasal in nature. Through a series of case studies, we undertake synchronic comparisons between academic writing and other spoken and written registers to demonstrate that this phrasal discourse style is highly distinctive – occurring only in written registers with highly informational purposes and specialized audiences. Thus, this chapter documents the distinctive grammatical characteristics of modern academic research writing by comparing it to conversation, textbooks, classroom teaching, fiction, and newspapers. We conclude Chapter 3 with a case study that considers variation *within* modern day academic research writing. Chapter 3 demonstrates that academic writing – and especially science research writing – can be regarded as an outlier: a register unlike nearly all other spoken and written registers.

Chapter 4 then turns to a diachronic analysis of this same range of registers, to document the evolutionary path that academic research writing has taken over the past 300 years. We investigate the use of colloquial features as well as complexity features, and we consider both clausal as well as phrasal complexity features. The results show that academic writing has been resistant to change in the use of colloquial

Table 1.4 *Challenging basic assumptions about grammatical complexity, linguistic change, and academic writing through corpus-based analyses*

	There is a single construct of grammatical complexity: structural elaboration realized through dependent clauses	Grammatical changes are initiated in speech; grammatical innovations do not occur in writing	All kinds of academic prose are essentially the same	Academic prose employs complex and elaborated grammar	Academic prose is maximally explicit in meaning	Academic prose is conservative and resistant to linguistic change
14	Chapter 3: Phrasal vs. clausal discourse styles (synchronic)	✓		✓		
	Chapter 4: The historical evolution of phrasal discourse styles	✓	✓	✓		✓
	Chapter 5: The functional extension of phrasal devices	✓	✓	✓	✓	✓
	Chapter 6: The loss of explicitness	✓		✓	✓	

features, but it has actually been the locus of grammatical change in the use of phrasal complexity features. Comparison of sub-registers within academic writing show that specialist science research articles have been especially important for these grammatical innovations.

Chapter 5 presents detailed linguistic analyses to support the claim that these historical developments are not ‘merely’ stylistic changes in frequency. Rather, this chapter documents major extensions in the lexical associations, meanings, and discourse functions of phrasal complexity features. Chapter 6 continues our functional analysis of the phrasal discourse style of academic writing, turning to an unintended consequence of these grammatical innovations: the loss of explicitness. Through detailed analyses of particular phrasal devices, we discuss how phrasal noun modifiers compress information into dense constructions that are considerably less explicit in meaning than alternative clausal features.

Taken together, the analyses presented in Chapters 3 through 6 directly challenge many stereotypes and assumptions about grammatical complexity, historical change, and academic writing. We summarize the overall progression of the book in Table 1.4.

The final chapter, Chapter 7, summarizes the main themes of the book, leading into a discussion of applied implications. First, the chapter revisits these basic assumptions about complexity, linguistic change, and academic writing, highlighting how the analyses presented in Chapters 3–6 counter previous stereotypes. We then turn to the implications that result from our corpus-based findings. We discuss implications for theories of grammatical complexity, arguing that phrasal modifiers and structural compression are equally important to, but fundamentally different from, clausal modifiers and structural elaboration. We also consider implications for linguistic change, arguing that grammatical change *can* emerge in writing, and that in fact certain types of linguistic change must be considered within the context of written language. Finally, we explore applied implications related to the study of language development and the teaching of academic reading and writing (English for Academic Purposes).

Using corpora to analyze grammatical change

2.1 Introduction

Historical research is observational and descriptive, based on consideration of grammatical features in historical texts in comparison to their use in present-day texts. Grammatical innovations (i.e., the emergence and development of new grammatical features) are readily observed by comparing texts from different historical periods. Thus, grammatical changes like the loss of case suffixes or the introduction of the dummy auxiliary DO are easy to notice in almost any sample of historical texts – even a relatively small sample.

In contrast, it is much more difficult to study shifts in the typical use and functions of grammatical features from a few selected texts. As noted in [Chapter 1](#), this type of grammatical change is more quantitative, involving expansions in frequency, range of lexical co-occurrence, and functionality. While it might be possible to notice a marked increase in frequency from a few texts, there is no way to study this range of phenomena systematically without access to a large and representative collection of texts: a corpus.

Many recent descriptions of grammatical change, like Krug (2000), Mair (2006), and Leech et al. (2009), have employed corpus-based analyses. One advantage of a corpus-based approach is its emphasis on the representativeness of the text sample, including deliberate sampling from different registers, and inclusion of as many different texts as possible. Thus, the first step for such analyses is to construct a corpus that represents the targeted language varieties and historical periods. For example, because Mair (2006) and Leech et al. (2009) were interested in twentieth-century change in general written English, comparing American English (AmE) to British English (BrE), they used the ‘Brown-family’ of corpora. These are one-million-word corpora designed to replicate the first large corpus of English: the Brown Corpus, which consists of 500 AmE written text samples published in 1961, taken from 15 text categories (e.g., newspaper reportage,

editorials, biographies, fiction, academic prose). Parallel corpora with this same design have been constructed for 1992 AmE (the Frown Corpus), 1961 BrE (the LOB Corpus), 1991 BrE (the F-LOB Corpus), 1931 BrE (the B-LOB Corpus), and 1901 BrE (the Lancaster BrE Corpus). Many other studies have used the ARCHER Corpus, a 1.8-million-word corpus of texts, organized in terms of eight speech-based and written registers sampled from 1650–1990 (see Biber, Finegan, and Atkinson 1994; Yáñez-Bouza 2011). And more recently, studies have been using the Corpus of Historical American English (COHA), a 400-million-word collection of language from fiction (c. 50% of the total), magazines, newspapers, and other books (available online at <http://corpus.byu.edu/coha/>; see Davies 2012). This corpus is structured in decades from 1810 to 2010, with a target of c. 20 million words sampled from each historical period (earlier periods have smaller samples; more recent periods are represented by larger samples).

As noted above, corpus-based methods are especially useful for the language-use perspective on grammatical change, analyzing quantitative patterns of change that proceed gradually over time. Corpus analysis facilitates three general types of analysis:

- 1) studying the use of core grammatical features in quantitative terms with a high degree of reliability, to document variation and change across historical periods and registers;
- 2) studying the choice between grammatical variants with considerable detail and precision, considering the relative influence of numerous interacting contextual and discourse factors;
- 3) studying lexico-grammatical change: the evolving grammatical use of a particular word, or the expansion of a grammatical innovation applied to an increasing set of words.

The first type of analysis is probably the most typical application of corpus analysis in historical research, documenting the course of linguistic changes in much more detail than previously possible. Such analyses have been undertaken to document shifts in the use of core grammatical features, such as the increase in the use of progressive verbs and semi-modal verbs, or the decrease in the use of modal verbs (see, e.g., Krug 2000; Biber 2004; Leech et al. 2009; Leech 2011). For example, Figure 2.1 plots historical change in the use of modal verbs in fiction and newspaper prose, based on analysis of the Corpus of Historical American English (COHA; Davies 2012). By basing such analyses on hundreds of texts, we are able to reliably document long-term historical trends, such as the gradual decline in the use of modal verbs shown in this figure.

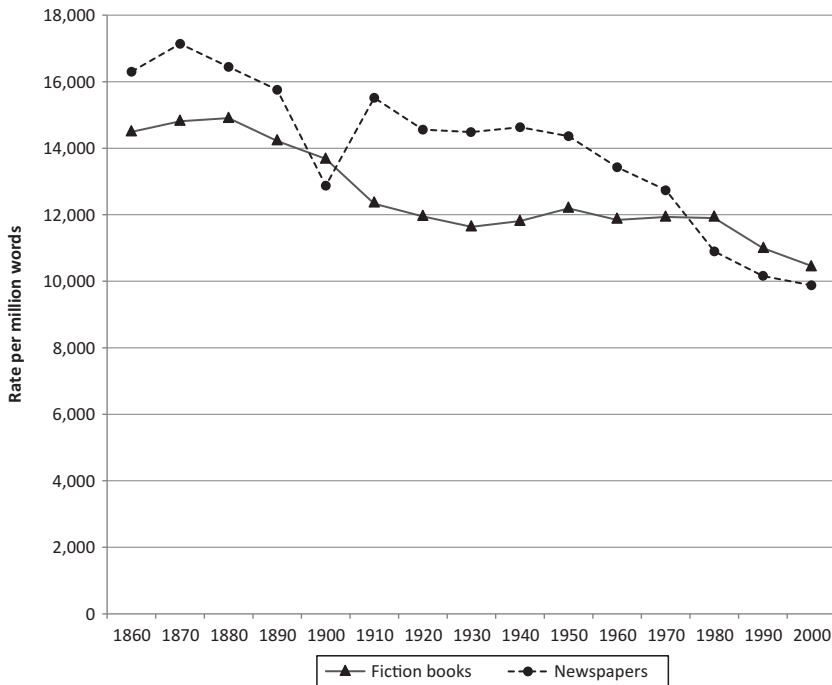


Figure 2.1 Decrease in the use of modal verbs in AmE fiction and newspapers (based on COHA)

It could be argued that the other two types of analysis were not even possible before the advent of historical corpora. In studies of grammatical variation, linguistic features with multiple structural realizations are investigated to determine the contextual factors that favor one variant over another, and how those preferences change over time. Example studies of this type include investigations of the choice between *of*-genitives and *s*-genitives (e.g., Hinrichs and Szendrői 2007), and the choice between ditransitive versus prepositional dative constructions (e.g., Wolk et al. 2013). In this case, the corpus provides thousands of instances of the grammatical construction that could have been interchangeable with another variant. By analyzing contextual factors such as the length of the phrase, information status, and animacy, it is possible to determine the factors that favor one or the other variant, and to document overall changes to the probabilistic grammar.

Finally, in the third type of study, a corpus is used to track the ways in which a grammatical construction comes to be used with an increasing set

of words, or the ways in which a word comes to be used with new grammatical functions (grammaticalization). For example, Leech et al. (2009: 129–130) investigate the claim that progressive aspect has come to be used with an increasing set of stative verbs in the twentieth century, based on analysis of the Brown family of corpora. Krug (2000) is a good example of a corpus-based grammaticalization study, describing historical change in the use of specific lexical combinations coming to be used as semi-modals (e.g., *have to*, *be going to*).

To illustrate, Figure 2.2 presents historical trends for *HAVE to* used as a semi-modal, based on analysis of COHA (Davies 2012). Because COHA is ‘tagged’ (i.e., words are coded for their grammatical characteristics), it is possible to identify occurrences that function as a semi-modal, by searching for *have/has/had* + *to* occurring immediately before a lexical verb. COHA does not cover periods earlier than 1810, and thus cannot be used

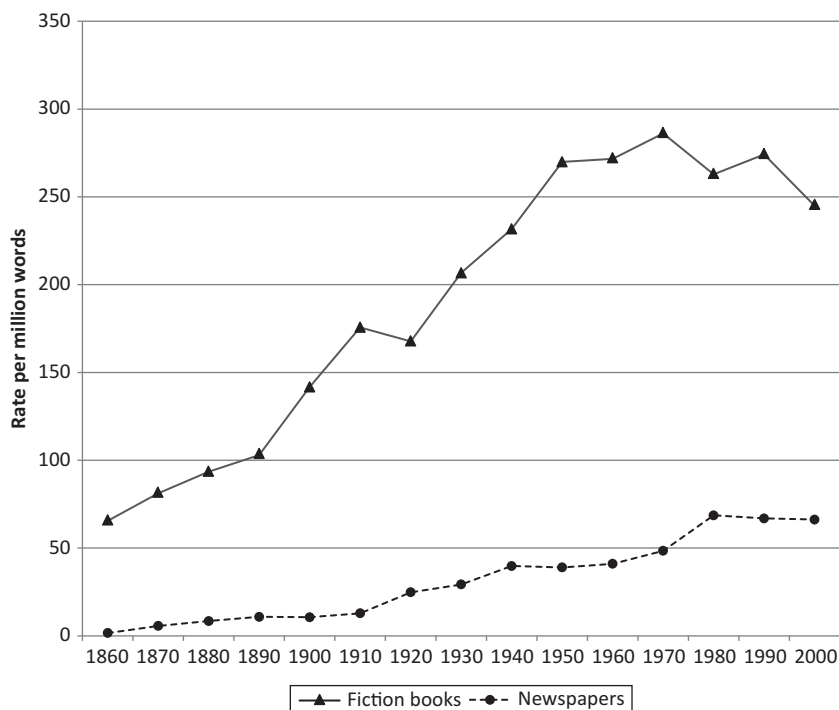


Figure 2.2 Increase in the use of the semi-modal *have to* in AmE fiction and newspapers (based on COHA)

to capture the earliest uses of *HAVE to* as a semi-modal. However, it does clearly represent occurrences from relatively early stages in this development, and also captures the very strong increase in frequency in later periods. Further, COHA allows us to isolate the historical developments within different registers, in this case showing a rapid increase in use in fiction, but a much smaller increase in newspapers.

A similar kind of grammatical innovation in English is the historical expansion in the number of multi-word verbs, including new phrasal verbs (e.g., *go ahead*, *get out*, *get NP back*, *run out*) and phrasal-prepositional verbs (e.g., *get away with*, *get back to*, *catch up with*). Phrasal verbs have existed for centuries in English. However, the lexical associations of this structural device have expanded considerably over the past two centuries, resulting in an increasing inventory of new phrasal verbs. For example, [Figure 2.3](#)

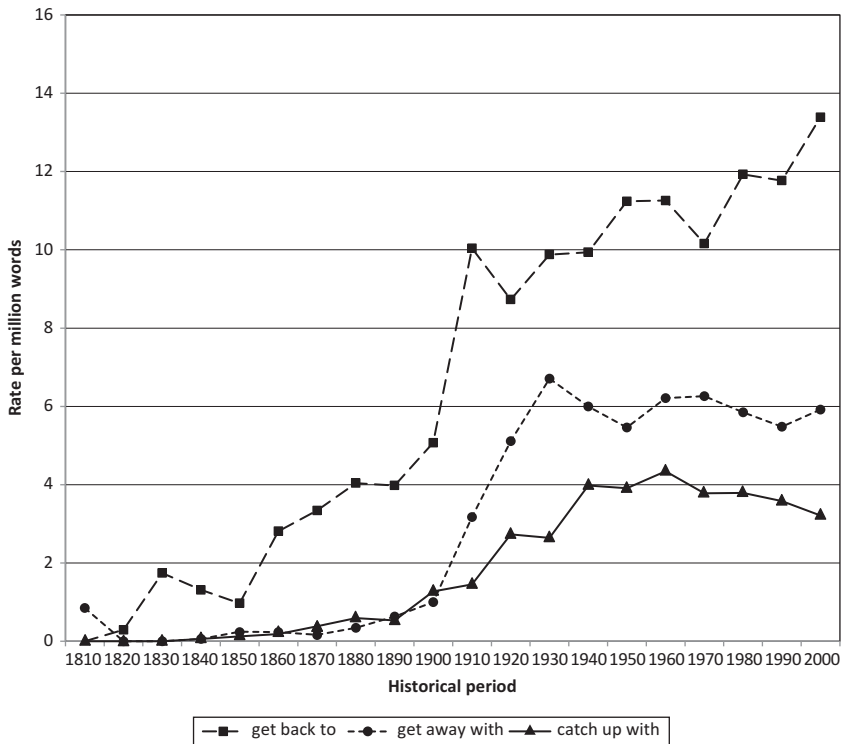


Figure 2.3 Increase in the use of specific phrasal-prepositional verbs in AmE (based on COHA)

shows the early use and subsequent increase for three new phrasal-prepositional verbs in COHA: *catch up with*, *get away with*, *get back to*. The overall trends here are clear, with a steady increase in use for all three verbs over the first half of the twentieth century. (The increase in use for *get back to* began a few decades earlier than the other two verbs.)

Such analyses require very large corpora, because lexical phenomena are much less common than grammatical phenomena, involving the use of particular words rather than the use of an entire construction. Thus, the actual numbers shown in Figure 2.3 are quite small, ranging from 4 to 10 occurrences per million words in the late twentieth century. To put this in perspective, imagine that a normal page of written prose is c. 500 words long. In that case, a one-million-word corpus would comprise 2,000 pages. And so a verb like *catch up with* would occur on only 4 of those 2,000 pages in a sample from the late twentieth century! The situation is actually even worse than it appears, since this rate of 4 times per million words is a kind of overall average, based on analysis of a much larger sample (e.g., c. 25 million words for the 1950 sub-corpus in COHA). Thus, it is likely that one one-million-word sample might contain more than 4 occurrences of *catch up with* – and that a different one-million-word sample might contain no occurrences at all!

Obviously, it would not be possible to document historical developments of this type without access to a large historical corpus. For example, imagine the frustration in reading thousands of pages of prose, in the hope of spotting a few occurrences of a specific multi-word verb. And then imagine repeating that process for each different verb, and for each different historical period. In general, historical descriptions of grammatical features can be based on analysis of smaller corpora, while lexical investigations require analyses of much larger corpora.

Thus, in the following chapters of the present book, we rely upon a range of corpora, including corpora that we have specially designed and constructed for our purposes, as well as established historical corpora that have been used in previous research (see Section 2.3).

It might be supposed that corpus-based research focuses exclusively on the quantitative analysis of general trends in a corpus. However, from our point of view, the best corpus-based research integrates quantitative and qualitative analyses, to document both changes in the *extent* to which features are used, along with *functional extensions* that develop over time. As Leech et al. (2009) point out:

It is also characteristic of corpus linguistics to begin with quantitative findings, and work towards qualitative ones. . . . Generally it is desirable

to subject quantitative results to qualitative scrutiny – attempting to explain why a particular frequency pattern occurs, for example. But on the other hand, qualitative analysis (making use of the investigator’s ability to interpret samples of language in context) may be the means for classifying examples in a particular corpus by their meanings; and this qualitative analysis may then be the input to a further quantitative analysis, one based on meaning. (Leech et al. 2009: 32)

Such paired qualitative and quantitative analyses are crucial for investigations of shifts in the typical functions and use of grammatical features over time and across registers. For example, once a representative, diachronic corpus is collected, it is a relatively straight-forward endeavor to document the quantitative increases in the frequency of grammatical structures like prepositional phrases as noun post-modifiers (e.g., *the increase in energy consumption*), and nouns as noun pre-modifiers (e.g., *oil prices*) (see discussion in Chapter 4). However, it is only through the complementary qualitative analyses of particular linguistic features that we can elaborate upon, and attempt to explain, the functional motivations for such quantitative trends of increased use. Without such complementary information, the identification of quantitative shifts remains simply that – a trend. In contrast, qualitative analysis allows us to discover the set of words typically used with a grammatical feature in different historical periods, the typical meanings expressed by those lexico-grammatical combinations, and historical shifts in the typical discourse functions served by the grammatical feature (see Chapter 5–6).

Leech et al. (2009) go on to highlight the importance of complementary *formal* interpretations (i.e., analyses of linguistic structures and associations between lexical and grammatical items) and situational interpretations (what Leech et al. call ‘functional’) of historical linguistic data:

On a formal level (thinking of grammatical analysis) we are interested in identifying *within* corpora particular distributional patterns of the occurrence (or co-occurrence) of structures and categories. On the functional level, we are interested in understanding and interpreting these patterns in terms of factors *external* to corpus data – cognitive, social or historical factors, for example. (Leech et al. 2009: 34)

Although these situational factors are not linguistic in nature, they are inherently linked to the context of language production. Such non-linguistic, or situational, factors are the foundation of a ‘register’ approach to language variation (see Biber and Conrad 2009), and corpus research studies have shown repeatedly that situationally-defined registers are strongly associated with variation in the linguistic characteristics of texts. Thus, the design of a historical corpus must take into account the

representation of historical periods as well as the representation of register differences within each period.

The primary focus of the present book is the kinds of grammatical innovations that have emerged in academic research writing. It might be supposed that the situational context of academic writing has not changed over the last four centuries. However, more careful consideration uncovers several social, situational, and technological patterns of change, including the rise of social science disciplines, increasing diversification among disciplines, increasing specialization within disciplines, and even the sheer amount of scientific inquiry taking place. We turn to a brief description of these historical developments in [Section 2.2](#), providing the basis for our corpus design in the present book.

2.2 Situational change within academic research writing

As we are interested in grammatical change in academic writing, it is important to consider the historical context of academic research (see Gross et al. 2002). Over the last 300 years, the world of academic research has undergone dramatic change, with the information explosion, the development and diversification of academic disciplines and specialized sub-disciplines, and the influence of technology on the role of research writing in society. In the seventeenth and eighteenth centuries, scholars were already publishing research studies in medicine and science. These were usually narrative reports of natural phenomena and processes that the author had observed. These researchers seemed to consider themselves to be scientists but not specialists in a particular discipline. Similarly, these research reports were read by the entire community of scientists (and even the wider literate public). The major disciplinary distinction of the time was between humanities academic writing (e.g., in philosophy, history, and literary discussions) and science/medical prose.

The nineteenth century was a period of transition in which scientific disciplines and specializations were beginning to form (e.g., biology, geology, astronomy, physics, chemistry). The twentieth century witnessed incredible disciplinary diversification within all of these general scientific disciplines. In addition, the general discipline of social science came to be recognized as a domain of inquiry, followed in short order by the development of numerous disciplines and sub-disciplines within social science. Science research during this period shifted from primarily observational to primarily experimental. By the end of the twentieth century, most science research articles were written by specialists in narrow sub-disciplines, to be

read by a small group of specialists from the same sub-disciplines. Over the course of the twentieth century, social science research also underwent a similar shift from observational to experimental designs, with a similar proliferation of sub-disciplines and research articles written for highly specialized audiences.

Over the course of the last three centuries, the genre of the science research article also underwent dramatic change (see Atkinson 1999; Gross et al. 2002; Biber and Conrad 2009: chapter 6). In the seventeenth and eighteenth centuries, many science articles were published as letters written to the editors of a journal. These texts were usually narrative reports, documenting what the author did and what they observed in the process. The nineteenth century witnessed a shift from observational to experimental research designs, and as a result, it became common for articles to include a section describing their research methods and procedures. This in turn led to the conventionalized form of science research articles in the twentieth century, with the major sections of Introduction-Methods-Results-Discussion. There is debate among discourse analysts about these historical changes, as to whether they represent change within a genre, or change to a new genre. We consider these both to be valid perspectives, capturing different aspects of the evolution of science research writing (see Biber and Conrad 2009: 165–166). However, this debate is only tangentially relevant to our primary goal in the present book: to document the patterns of historical change in the use of grammatical features in science research writing (compared to the grammatical styles employed in a range of other spoken and written registers). For that goal, it is essential that we consider the full range of register variation within each historical period, and then compare the patterns of use across periods – regardless of whether those texts are taken from the ‘same’ genre or different genres.

2.3 Corpora analyzed in the book

The social history of academic research writing, coupled with our own research goals, shaped the design of the corpora that we constructed for this project. Specifically, we assembled a set of corpora that enabled investigation of the following issues:

1. What grammatical features are typically employed in modern academic writing? How do those patterns of use compare to grammatical styles typically used in other spoken and written registers?

2. Are there systematic differences in the grammatical features commonly used in modern texts from different academic disciplines?
3. Are there systematic differences in the grammatical features commonly used in specialist versus generalist science research articles?
4. Have the synchronic grammatical patterns documented in 1–3 changed over time? If so, how? In particular, can we document the patterns of change for the distinctive grammatical features used in science research writing?
5. Have there been concomitant changes in the lexical associations, typical meanings, or typical discourse functions of these grammatical devices over time?

Because our primary focus is on professional academic writing, most sub-corpora analyzed in the following chapters are sampled from academic research journals, with some supplementation from academic books. These corpora, used for both synchronic and diachronic analyses, are described in [Section 2.3.1](#). However, because we can more readily observe linguistic change in academic writing by comparing it to other registers, we also undertake analyses of sub-corpora representing other written registers (fiction, newspapers, university textbooks) as well as spoken registers (e.g., conversation, dramatic dialogues, classroom teaching); these comparison sub-corpora are described in [Section 2.3.2](#).

2.3.1 Representing academic writing over time: Description of the diachronic academic corpora

The analyses that we undertake in this book are based on a collection of corpora that represent academic writing from 1700 to 2005, encompassing a range of sub-registers within academic writing. While it is relatively easy to obtain large historical samples for novels and non-fiction books, it is more difficult to obtain online academic research articles from earlier historical periods. For this reason, the corpus used for our project was assembled from various existing corpora. The resulting corpus is not carefully balanced for the number of texts occurring in each period and/or each sub-register. However, by combining the resources of other available corpora, we were able to assemble a diachronic corpus that is much more representative of scientific research writing than any pre-existing corpus would have been on its own.

Our original focus in the project was on grammatical developments that occurred in academic writing in the last half of the twentieth century

Table 2.1 *Academic research journals sampled for the 20th Century Research Article Corpus*

Science:	<i>Journal of Cell Biology, Biometrics, American Journal of Medicine, Journal of Animal Ecology, Journal of Physiology</i>
Social Science:	<i>American Educational Research Journal, Journal of Educational Measurement, American Journal of Psychology, Developmental Psychology</i>
Humanities:	<i>Journal of Contemporary History, Journal of the History of Ideas</i>

(see, e.g., Biber and Gray 2010). For those studies, we compiled the 20th Century Research Article Corpus, which contains published research articles from three twenty-year intervals (1965, 1985, 2005) in three general disciplines: science (medicine, physiology, biology, ecology), social sciences (education, psychology), and humanities (history). These disciplines were selected to represent a wide range of research within academia, from hard science to social science to humanities. Articles were randomly selected from eleven different research journals from the three major disciplines, for each of the three sampling years. Table 2.1 lists the research journals sampled for each of the three major disciplines.

The resulting corpus comprises 570 texts and c. 3.6 million words: 540 research articles, and 30 samples from 2005 research monographs in history. For some of the synchronic analyses reported in this book (in Chapter 3), results are reported for this corpus as a whole. However, we are also interested in exploring the extent to which relatively subtle sub-register and disciplinary differences *within* academic writing are influential as determinants of linguistic change. Thus, as Table 2.2 shows, the 20th Century Research Article Corpus can also be segmented into sub-corpora representing specialist science writing, specialist social science writing, and specialist humanities (history) writing. In addition, we include a sub-corpus of recent non-specialist multi-disciplinary science articles, published in either *The Philosophical Transactions of the Royal Society* or in the journal *Science* (see discussion below).

The 20th Century Research Article Corpus enables analyses of modern day academic writing and recent historical change in the last half of the twentieth century. However, to track longer-term historical change, we assembled sub-corpora from other existing corpora of academic writing. The complete corpus used for these longer-term analyses is summarized in

Table 2.2 *Design and composition of the 20th Century Research Article Corpus*

Sub-corpus	Year			Totals	
	1965	1985	2005	Texts	Words
Specialist science (biology, medicine, ecology, physiology research articles)	71	44	52	167 texts	c. 1.1 million
Specialist social science (education, psychology research articles)	58	58	57	173 texts	c. 1.1 million
Specialist humanities (history books and research articles)	30	30	60	120 texts	c. 930,000
Non-specialist (multi-disciplinary) science (<i>Philosophical Transactions, Science</i>)	10	60	40	110 texts	c. 520,000
Total	169 texts c. 930,000 words	192 texts c. 1.2 million words	209 texts c. 1.5 million words	570 texts	c. 3.6 million words

Table 2.3, and includes the 20th Century Research Article Corpus as well as samples for earlier historical periods obtained from multiple sources.¹

The primary corpus used for this purpose was ARCHER (*A Representative Corpus of Historical English Registers*; see Biber, Finegan and Atkinson 1994; Yáñez-Bouza 2011). ARCHER was originally designed for diachronic investigations of historical change in a wide range of written and speech-based registers from 1650 to present, with systematic 50-year samples across ten major register categories. While the complete corpus includes 1,037 text samples and c. 1.8 million words, for the purposes of this book, we relied on the academic sub-corpus of ARCHER, which includes texts in science and medicine (150 texts and c. 430,000 words). A second specialized science corpus was used to supplement our samples of science articles from the

¹ Most of the analyses reported in this book begin in 1700 and continue to the present day; thus, Table 2.3 begins at 1700. However, for a few analyses, we report corpus findings for 1650/seventeenth century based on analysis of *PT* or medical texts from that period.

Table 2.3 *Complete academic writing corpus used for long-term historical analyses in the book (number of texts)*

Register	18 th C.		19 th C.		20 th C.					Total
	1700	1750	1800	1850	1900	1925	1965	1985	2005	
Specialist science (biology, medicine, ecology, physiology research articles)							71	44	52	167 texts c. 1.1 million words
Specialist social science (education, psychology research articles)							58	58	57	173 texts c. 1.1 million words
Specialist humanities (history books/monographs/research articles)				37			30	30	60	157 texts c. 2.5 million words
Non-specialist (multi-disciplinary) science (<i>Philosophical Transactions</i> , <i>Science</i> , <i>ARCHER</i> , <i>CETA</i>)	30	31	50	101	50	50	10	60	40	422 texts c. 2.4 million words

eighteenth and nineteenth centuries: the *Corpus of English Texts on Astronomy* (CETA; see Crespo García and Moskowich-Spiegel Fandiño 2010 for a full description). The sample used for our project comprised 42 astronomy texts and c. 400,000 words.

Our sub-corpus of science research articles for the nineteenth century was supplemented by samples from the *Philosophical Transactions of the Royal Society* (PT, published since 1665) and *Science* (published since 1880). For PT, we collected samples of twenty texts from 1800, 1850, 1875, 1900, and 1925; for *Science*, we collected samples of twenty texts from 1880, 1900, and 1925. Pilot research for the project had indicated that the nineteenth and twentieth centuries were especially important historical periods of development for the grammatical features focused on in our study. In our earlier research, we had already developed a relatively large corpus of research articles from the twentieth century, and this additional inclusion of nineteenth century articles from PT and *Science* provided a more solid foundation for our analyses of historical change in that period.

Our pilot research also uncovered an additional unanticipated finding that shaped our corpus design. Previous corpus-based research on the history of science writing had relied on PT, because it had a continuous history of publication since 1665. Thus, our initial inclination was to use this same journal as the primary source of science research articles in our study. By restricting the analysis to articles published in a single journal, we expected that we would remove possible unwanted variation associated with different journals. However, we discovered in our pilot research that the grammatical characteristics of PT articles in the late twentieth century were systematically different from the grammatical characteristics of specialist science research articles published in other journals. More careful consideration showed that PT had to a large extent remained relatively constant in purpose and readership over the centuries, intended as outlets for the most important science research findings from across the full range of scientific disciplines, and written for an audience from that same breadth of disciplines. In the nineteenth century, most science writing was published in multi-disciplinary venues and read by a wide multi-disciplinary audience. Thus, journals like PT provide a fairly good representation of the universe of science writing from that period. In contrast, articles of that type represent only a very small proportion of modern science writing. Instead, most science writing today is published in highly specialized journals associated with specific sub-disciplines, and read mostly by specialists from those same sub-disciplines. It turns out that these differing goals and intended audiences are associated with important grammatical differences. Thus, we added extra samples of

texts from *PT* (as well as the journal *Science*) for the late twentieth century, allowing us to directly compare the historical patterns of change in generalist multi-disciplinary science writing with those found in specialist science research articles. Specifically, we added samples of twenty texts per period for articles published in *PT* (in 1985 and 2005), and twenty texts per period for articles published in *Science* (in 1985 and 2005).

Finally, to enable longer-term comparisons for humanities writing, a sub-corpus of history monographs (37 texts, c. 1.6 million words) was collected for 1850–1900, supplementing the sub-corpus of history research articles and books from the late twentieth century.

2.3.2 *Synchronic and diachronic comparisons between academic writing and other spoken and written registers*

The distinctive grammatical characteristics of academic writing are best appreciated when they are described in comparison/contrast to other spoken and written registers. We used a range of corpora for these comparisons, as described in Table 2.4. Present-day corpora for face-to-face conversation, newspaper articles, and modern fiction are taken from the *Longman Spoken and Written English (LSWE) Corpus* (see Biber et al. 1999: 29–35 for a full description). Constructed for the corpus-based analyses in the *Longman Grammar of Spoken and Written English*, these corpora include both British and American English and are widely regarded as large, representative samples of modern day registers.

In addition, sub-corpora from the *TOEFL-2000 Spoken and Written Academic Language (T2K-SWAL) Corpus* have been used to compare academic writing to two other academic registers: academic textbooks (87 text samples, c. 760,000 words) and classroom teaching (176 class sessions, c. 1.2 million words). For a full description of this corpus, see Biber (2006a, Chapter 2).

Diachronic comparison corpora were constructed for three general registers: newspaper prose, fiction, and speech-based dialogue (represented by drama). The historical fiction sub-corpus is taken from ARCHER (99 text samples published between 1700–1990), supplemented with 116 full-text novels (1700–1990) collected from *Project Gutenberg* (www.gutenberg.org/). The historical newspaper corpus was also taken from ARCHER (90 texts, 1700–1990), supplemented with 250 articles published in *The New York Times* from 1850 to 2000 (50 texts published in 1850, 1900, 1950, and 2000). The historical magazine corpus includes 800 articles from the twentieth century (1925, 1965, 1985, 2005), all taken from the ‘National Affairs’ or ‘Foreign News’ sections of *Time Magazine*. The historical

Table 2.4 *Description of comparison corpora*

Synchronic Comparisons			
	<i>Texts</i>	<i>Words</i>	
Conversation (Longman Corpus of Spoken and Written English, Biber et al 1999)	723	c. 6.4 million	
Newspapers (Longman Corpus of Spoken and Written English, Biber et al 1999)	c. 32,000	c. 10.6 million	
Fiction (Longman Corpus of Spoken and Written English, Biber et al 1999)	139	c. 5 million	
Academic Textbooks (T2K-SWAL Textbooks sub-corpus, Biber 2006a)	87	c. 760,000	
Classroom Teaching (T2K-SWAL Classroom Teaching sub-corpus, Biber 2006a)	176	c. 1.2 million	
Diachronic Comparisons			
	<i>Texts</i>	<i>Words</i>	<i>Time Period</i>
Historical Fiction (ARCHER, <i>Gutenberg</i> novels)	215	1.2 million	1700–1990
Historical Newspapers (ARCHER, <i>New York Times</i>)	340	250,000	1700–1990
Historical Magazines (<i>Time Magazine</i>)	800	560,000	1925–2005
Historical Drama (ARCHER)	85	c. 75,000	1700–1990

newspaper and magazine corpora were combined for the purposes of the correlational analyses in [Chapter 4](#).

2.4 Corpus analysis procedures

All corpora were grammatically annotated using the Biber ‘tagger’. This software package was originally developed in the early 1990s, and then greatly expanded as the basis of the analyses reported in the *Longman Grammar of Spoken and Written English* (Biber et al. 1999). The Biber

tagger relies upon probabilistic and rule-based algorithms to annotate each word in a text for detailed part-of-speech and syntactic information. Several specialized programs were subsequently utilized to process these tagged texts, identifying each instance of a range of grammatical features, and computing rates of occurrence normalized to 1,000 words (see Biber, Conrad, and Reppen 1998: 263–64).

Rates of occurrence were computed for each individual text in the corpora. That is, each ‘text’ is considered the unit of observation for all analyses. This research design is atypical of most quantitative corpus research, which tends to instead report a single frequency count for the entire sub-corpus. By calculating rates of occurrence for each text (rather than a single count for each sub-corpus), it is possible to compute a genuine mean score for each register and a standard deviation (measuring the extent to which linguistic scores vary across texts within a register). With this research design, it is also possible to employ inferential statistical tests to determine whether the differences among registers are statistically ‘significant’ and to statistically measure the ‘strength’ of observed differences among registers. For example, we are able to compute Pearson correlations between the linguistic rates of occurrence and the year of the text.² These correlations measure linear historical trends, regardless of the overall extent to which a feature is used. Likewise, we employ various ANOVA analyses to measure the degree to which mean differences across registers and across historical periods are statistically different from one another. None of those analyses are possible with the traditional design based solely on the frequency of a feature in the sub-corpus. (See the extended discussion in Biber and Jones 2009; Biber 2012.)

2.5 Grammatical features investigated

We investigated a wide array of grammatical features in the study, ranging from core grammatical features (e.g., nouns, adjectives, verbs, adverbs) to highly specialized constructions (e.g., noun-participle compounds functioning as noun pre-modifiers). In the present section, we introduce these features, providing the foundation for the corpus analyses in later chapters. We group features into three sets: ‘core’ grammatical features; ‘colloquial’ grammatical features; and features associated with clausal and phrasal complexity.

² Pearson correlations measure the extent to which two variables have a linear relationship. However, historical change often progresses in a non-linear manner, in which case the actual (non-linear) correlations would be higher than those reported here. A non-parametric correlation like Spearman is more conservative regarding assumptions of linearity, but as a result inflates the estimate of the relationship. Thus, we use Pearson correlations as a more conservative measure.

The following sections give a general structural introduction to these features, which are then described from a language-use perspective in [Chapters 3–6](#). While most of these features can be identified automatically in a corpus using specialized computer programs (described in [Sections 2.4](#) and [2.5.3](#)), some of them require hand-coding (e.g., prepositional phrases as nominal post-modifiers). In the present section, we identify the features that required such interactive coding, and provide details regarding the specific samples that were hand-coded.

2.5.1 *‘Core’ grammatical features and ‘colloquial’ features*

Although the primary focus of our study is on phrasal complexity features in academic writing, we also analyze the use of two other groups of features. To document the general characteristics of academic writing compared to other registers in English, we provide quantitative data on several ‘core’ grammatical features: the content word classes (nouns, lexical verbs, adjectives, and adverbs), plus two grammatical categories associated with academic writing (passive voice verbs and nominalizations). All features in this group were identified and counted automatically using grammatical tags.

A second set of features are associated with conversational interaction. These are relevant to our study because several previous researchers have claimed that written registers have become increasingly ‘colloquial’ in recent decades, and thus they use these conversational features to a greater extent than in previous historical periods. In [Chapter 4](#), we focus on the use of fifteen features ([Table 2.5](#); [Table 4.2](#)) that are especially common in modern-day conversation (based on a survey of the corpus findings in the *Longman Grammar of Spoken and Written English*, Biber et al. 1999); these features include many of the most important colloquial features that have been discussed in previous research on the colloquialization of written language.

2.5.2 *Phrasal and clausal complexity features*

Grammatical complexity is linked with structural elaboration and the embedding of clauses and phrases in linguistic theory. A ‘simple’ clause has only a subject, verb, and object or complement. A ‘simple’ noun phrase has a determiner and head noun. Structural additions to simple clauses or noun phrases represent elaboration, resulting in ‘complex’ grammar.

There are many different types of grammatical features that can be used for structural elaboration, including embedded finite clauses, non-finite clauses, and phrases. However, previous corpus research has also shown that

Table 2.5 *Colloquial grammatical features investigated in the study*

Feature	Examples
Contractions	<i>can't, wouldn't, she's, won't</i>
Core Modals	<i>may, might, should, would, could</i>
Semi-modals	<i>have to, be going to, got to</i>
Progressive Aspect	
simple progressive	<i>is living, am going, was saying</i>
perfect progressive	<i>has been keeping, had been waiting, have been talking</i>
Other Verb Features	
GET-passive	<i>got left, get involved, get stuck</i>
help + bare infinitive	<i>help keep things together, help explain why. . .</i>
phrasal verbs	<i>come on, pick up, look up, turn out</i>
Personal Pronouns	
1 st person pronouns	<i>I, we</i>
2 nd person pronouns	<i>you</i>
3 rd person pronouns	<i>he, she, they</i>

Table 2.6 *Examples of complexity features, by structural type versus syntactic function*

Parameter A: Grammatical Type	Parameter B: Syntactic Function	
	Clause Constituent	Noun Phrase Modifier
<i>Finite Dependent Clause</i>	conditional clause	WH relative clause
<i>Non-finite Dependent Clause</i>	to-verb complement clause	non-finite passive relative clause
<i>Dependent Phrase</i>	adverb phrase as adverbial	prepositional phrase as noun modifier

syntactic function is equally important to structural type for describing grammatical complexity in spoken versus written registers. For example, spoken registers tend to utilize embedded structures that function as clause constituents (adverbials or verb complements), while written registers tend to utilize embedded structures that function as noun phrase modifiers.

In the present study, we tried to include the full suite of grammatical features that could be associated with grammatical complexity. These features can be grouped according to the two parameters of structural type versus syntactic function. Table 2.6 gives example grammatical features for each category. For example, a WH relative clause is a 'finite dependent clause' along *Parameter A: Grammatical Type*, which functions syntactically as a modifier within a noun phrase (*Parameter B: Syntactic Function*).

Thus, the full set of grammatical complexity features analyzed in the following chapters are presented in Table 2.7. In the table, grammatical features are grouped primarily by *Parameter A: Structural Type*, and secondarily by *Parameter B: Syntactic Function*.

Finite dependent clauses can function as adverbials (causative, conditional, or concessive clauses), complements (verb complements, adjective complements, or noun complements), or noun modifiers (relative clauses). Non-finite dependent clauses can similarly function as adverbials (purpose clauses), complements (verb complements, adjective complements, or noun complements), or noun modifiers (non-finite passive or *-ing* relative clauses). Finally, dependent phrases are realized as many different specific structural types (e.g., noun phrases, adjective phrases, prepositional phrases), functioning as clause modifiers (i.e., adverbials) or as noun-phrase modifiers. These features have been less commonly analyzed in previous studies of grammatical complexity.

Both parameters (structural type and syntactic function) turn out to be important for understanding the grammatical complexity of written academic discourse. We have employed this framework in many recent analyses, including comparisons of: spoken versus written registers (e.g., Biber and Gray 2010, 2013a; Biber, Gray and Poonpon 2011, 2013); sub-registers within academic writing (e.g. Biber and Gray 2010, 2013b; Gray 2015b); academic disciplines (Gray 2015a, 2015b); and even proficiency levels in novice academic writing (Biber, Gray, and Staples 2014; Staples et al. in press).

The two parameters can be considered together as a single cline of grammatical complexity, illustrated in Figure 2.4.

The analyses presented in the following chapters show that the features of phrasal complexity are especially innovative: these features have changed the most in terms of frequency and discourse function over the last few centuries. These complexity features are ‘phrasal’ with respect to both parameters: they are phrasal structures, and they function as noun phrase modifiers. It turns out that the grammatical discourse style characterized

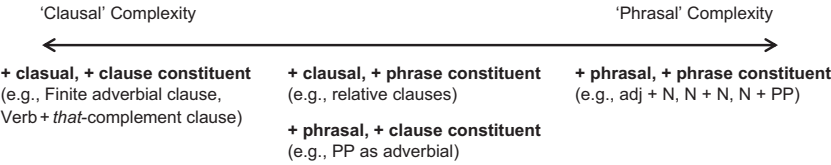


Figure 2.4 A cline of grammatical complexity

Table 2.7 *Complexity features investigated in the study, by structural type versus syntactic function*

Parameter A: Structural Type	Parameter B: Syntactic Function	Grammatical Sub-Categories	Examples
1. Finite Dependent Clauses	1A. Adverbial	Causative clauses: <i>because</i> + <i>clause</i>	<i>She won't narc on me, <u>because she prides herself on being a gangster.</u></i>
		Conditional clauses: <i>if</i> + <i>clause</i>	<i>Well, <u>if I stay here,</u> I'll have to leave early in the morning.</i>
		Concessive clauses: <i>although</i> + <i>clause</i>	<i>If I don't put my name, she doesn't know who wrote it, <u>although she might guess.</u></i>
	1B. Complement	Verb controlled <i>that</i> -clause	<i>I would hope <u>that we can have more control over them.</u></i> (with ZERO complementizer): <i>yeah, I think <u>I probably could.</u></i>
		Verb controlled <i>wh</i> -clause	<i>I don't know <u>how they do it.</u></i>
		Adjective controlled <i>that</i> -clause	<i>It is evident <u>that the virus formation is related to the cytoplasmic inclusions.</u></i>
		Noun controlled <i>that</i> -clause	<i>The fact <u>that no tracer particles were found in or below the tight junction (zonula occludens) indicates that these areas are not a pathway for particles of this size in the toad bladder.</u></i>
	1C. Noun modifier	Relative clause with <i>that</i>	<i>The results from a large number of cloze tests were used to estimate the amount of experimental error <u>that could be expected to result from using cloze tests of various lengths</u></i>
		Relative clause with <i>wh</i> -relativizer	<i>Their nucleoid is formed by dense granules and rods composing a ring <u>which limits a central electrontransparent space.</u></i>

Table 2.7 (*cont.*)

Paramater A: Structural Type	Parameter B: Syntactic Function	Grammatical Sub-Categories	Examples
2. Non-finite Dependent Clauses	2A. Adverbial	<i>to</i> -clause indicating ‘purpose’	<i><u>To verify our conclusion that the organic material is arranged as a coating around the silica shell components, thin sections of fixed cells were also examined.</u></i>
	2B. Complement	Verb-controlled <i>to</i> -clause	<i>I really want <u>to fix this room up.</u></i>
		Verb-controlled <i>ing</i> -clause	<i>I like <u>watching the traffic go by.</u></i>
		Adjective-controlled <i>to</i> -clause	<i>It was important <u>to obtain customer feedback.</u></i>
		Noun-controlled <i>to</i> -clause	<i>The project is part of a massive plan <u>to complete the section of road.</u> . .</i>
	2C. Noun modifier	Noun + <i>ing</i> -clause (non-finite <i>-ing</i> relative clause)	<i>Transfer tests <u>following over-training</u> indicated individual variability.</i>
Noun + <i>ed</i> -clause (non-finite passive relative clause)		<i>The results <u>shown in Tables IV and V</u> add to the picture. . .</i>	
3. Dependent phrases (non-clausal)	3A. Adverbial	Adverb phrase	<i>I raved about it <u>afterwards.</u></i>
		Prepositional phrase	<i>Alright, we’ll talk to you <u>in the morning.</u></i>
	3B. Noun Modifier	Attributive adjectives as noun pre-modifiers	<i><u>emotional injury</u>, <u>conventional practices</u></i>
		Nouns as noun pre-modifiers	<i><u>the trial transfer sessions</u></i>
		Prepositional phrases as noun post-modifiers	<i>Class mean scores were computed by averaging the <u>scores for male and female target students in the class.</u></i>
		Appositive noun phrases as noun post-modifiers	<i>Two Stuart monarchs (<u>Charles I and Charles II</u>) were strongly suspected of Romish sympathies. James Klein, <u>president of the American Benefits Council</u></i>

by the dense use of these features is especially characteristic of modern science research writing.

2.5.3 Corpus analyses of grammatical complexity features

As mentioned above, all corpora were tagged using the Biber tagger and then processed using specialized computer programs to generate per-text rates of occurrence for a range of features. These more specific programs incorporated lexico-grammatical information from *The Longman Grammar of Spoken and Written English* (Biber et al. 1999), such as lists of common nouns that can control a *that*-complement clause (e.g., *fact*, *hypothesis*, *claim*) or a *to*-complement clause (e.g., *proposal*, *plan*, *bid*). This approach allowed us to analyze syntactic features that could otherwise not be identified reliably using automatic techniques. The results of these analyses were checked by hand to ensure accurate identification of the target features.

While many of the features listed in Table 2.7 could be identified and counted automatically, other features required hand-coding. In most cases, rates of occurrence for such features are based on a hand-coding of a subsample of the corpus data, which are then used to arrive at adjusted estimates of their use. We note details of those procedures in Chapters 3–6, where we discuss the results of those features. The paragraphs below outline the general approach taken for these analyses.

Prepositional phrases are especially important for our purposes here. Prepositional phrases can function as either adverbials or as noun modifiers, but automatic computer programs have limited success in determining the function of many occurrences. Thus, consider the following two example sentences. Structurally, there is no overt indication of syntactic role in these examples: both examples have an *in*-phrase immediately following a noun. But in example (a), the underlined prepositional phrase functions as a noun modifier, while the underlined prepositional phrase in example (b) functions as an adverbial:

- a) Some of the particles in the membrane are indicated by the brackets.
- b) The Concept of Dread again takes up the question in the section to which reference has already been made.

This distinction is important for our analyses in the present book, because prepositional phrases functioning as noun modifiers are pervasive in academic writing, while prepositional phrases functioning as adverbials are common in many different registers. Nouns followed by *of*-phrases could

be analyzed automatically because they almost always function as post-nominal modifiers. Similarly, prepositions preceded by a verb are clearly either adverbials or part of a phrasal verb complex. Other prepositions require human coding to determine the syntactic function of the prepositional phrase. Thus, we designed a computer program that would select a random sample of occurrences for particular prepositions (*in*, *on*, *with*, and *for*), outputting concordance lines for every *n*th instance of the preposition when it was not preceded by a verb. These concordance files were then manually coded to differentiate between prepositional phrases functioning as a noun modifier versus those functioning in all other roles (e.g., adverbials).

Noun phrases were also coded by hand to identify appositive noun phrases that function as noun post-modifiers. For that analysis, we developed a computer program to automatically identify all occurrences of grammatical sequences that might be appositive noun phrases:

NP + COMMA + NP e.g., James Klein, president of the American Benefits Council

NP + PARENTHESIS + NP e.g., the other member of each twin pair (twin B sample); Protein expression was assessed by Western blot (anti-myosin tail antiserum)

All of these occurrences were then manually coded to determine if they were appositive noun phrases versus instances of some grammatical structure.

*Phrasal versus clausal discourse styles: A synchronic
grammatical description of academic writing
contrasted with other registers*

3.1 Introduction

Academic writing is one of the most studied registers of English. There are several reasons for all this attention. The most cynical explanation is that academic scholars are simply narcissistic. We study academic prose because it is the register that we are most familiar with, we admire it because it is what we do best, and so we find it to be inherently important and worthwhile.

But academic writing is also important because of its role in university education: it is the primary register that students must control for academic success. There is no doubt that the challenges of learning to read and write academic prose is one of the major hurdles of advanced education. As a result, applied linguists have invested considerable energy in the study of this register, providing a descriptive foundation for reading and writing instructors trying to teach the intricacies of academic writing to novice students.

Beyond that applied motivation, there is a more important reason for the attention given to academic writing: it is dramatically different linguistically from spoken discourse, and actually different from most other written registers as well. This linguistic difference is what makes academic writing so challenging for students. And this distinctiveness makes academic writing worthy of investigation, to document its linguistic characteristics and track its historical development.

The distinctive linguistic characteristics of any register can be attributed to the situational context of that register, and academic prose is no exception. For example, academic writing is influenced by the circumstances under which it is produced; academic writers can take as much time as they need to plan exactly what they want to write, and if they write something unintended, they can delete/add/revise/edit the language of the text. Thus, the final written text that an external reader sees might bear

little resemblance to the initial words that the author produced, and readers usually have no overt indication of the extent to which the author has revised the original draft.

And production circumstances are only one of the ways that academic writing differs from other registers. For example, academic writing differs from spoken conversation in just about every conceivable way. Conversation is addressed to specific individual participants, it is highly interactive, the speaker and hearer are present together in the same situation, speakers discuss personal topics, and it is produced in real time with no pre-planning. In contrast, academic writing is addressed to a large readership which does not normally interact with the author. Individuals from that group can read the academic text in any location at any time. Although the author does not know members of this group individually, there is a high degree of shared background knowledge, especially in the case of research articles written for specialists in an academic discipline. Academic prose always has informational purposes, but it can also be overtly persuasive to differing extents.

At the same time, ‘academic writing’ itself can be said to encompass several related written registers, each with its own set of situational characteristics. To illustrate, consider two of the situational parameters that distinguish academic writing from non-academic registers: informational versus non-informational purpose, and specialist versus non-specialist audience. As Figure 3.1 shows, these same two parameters also differentiate among specific sub-registers within ‘academic writing’ (marked in *italics* in the figure).

Figure 3.1 illustrates how registers vary to differing extents with respect to situational parameters like purpose and audience. It is not the case that a text or register is simply informational (or not), or that the audience is simply specialist (or not). Rather, registers differ in the extent to which they have informational purposes and specialized audiences. For example, newspapers and non-fiction books are both informational in purpose (their goal is to provide facts, accounts of events, etc.), and written for general (non-specialist) audiences with little assumed background knowledge. But the two registers also differ from each other: non-fiction books can be regarded as more informational in purpose, and written for a more specialized audience, than newspapers.

Thus consider the popular non-fiction work *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything* (Levitt and Dubner 2009), which takes the results from primary economics research, applies them to problems in everyday life, and discusses them in a humorous account. The book includes extended explanations of economic principles,

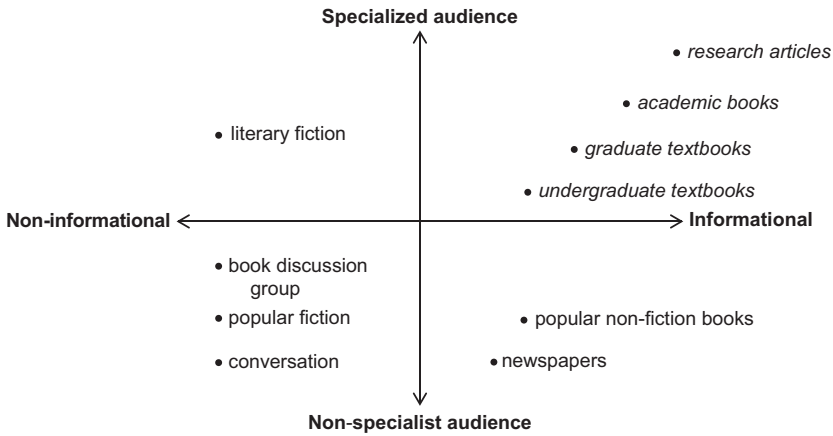


Figure 3.1 Academic and non-academic registers along two situational parameters

packaged for readers who demonstrate enough interest in the economics of everyday life to read a book-length treatment on the topic. This book can be contrasted with the newspaper article “Stung last year, retailers and shippers retool for the holiday season” (by Rachel Abrams), published in the *New York Times* (Nov. 3, 2014). That article is also about economics, but the information and intended readership is much more general, concerning the upcoming holiday shopping season and what retailers are doing to prepare, written to inform any reader interested in current events.

While popular non-fiction books and newspapers share an informational purpose with academic writing, other registers are further from academic writing with respect to purpose. For example, a popular fiction book like Dan Brown’s *The DaVinci Code* has the purpose of providing entertainment for millions of readers around the world, none of whom are expected to have any degree of background knowledge about symbology or cryptology (two of the main themes that drive the book’s plot). In contrast, a book like Flannery O’Connor’s *Everything that Rises Must Converge* can be regarded as literary fiction, written for non-informational purposes (with a focus on character and theme development), with a much more specialized readership than popular fiction novels.

Figure 3.1 illustrates how written academic registers are distinctive when compared to most other registers, having a primary informational purpose and being written for audiences with varying degrees of specialist knowledge. However, Figure 3.1 also captures differences in the situational characteristics of different sub-types of academic writing. For example,

undergraduate textbooks such as David Myers' *Psychology* (10th Edition, 2011) are intended to provide students in an introductory psychology class with well-established knowledge across a range of topics in psychology (e.g., aging, anxiety, depth perception, eating disorders, memory, brain abnormalities).

Graduate-level textbooks are increasingly informational and targeted towards students with a greater degree of background knowledge about psychology. Thus, a text like Wampold's (2001) *The Great Psychotherapy Debate: Models, Methods, and Findings* will present more detailed information on a narrower topic, which requires a previous knowledge of psychology in general and psychotherapy more specifically.

Continuing towards the upper right-hand corner of Figure 3.1, research articles are written for a highly specialized audience and focused almost entirely on presenting conceptually-demanding information. The authors of research articles work to establish new knowledge, often challenging previous claims. The readers of these articles are also specialists in a particular topic or sub-discipline, and perhaps even conducting research in the same tradition. Thus, a psychology research article such as 'Mediation analysis allowing for exposure-mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros' (Valeri and VanderWeele 2013, *Psychological Methods* 18(2), 137–150) is written for researchers working in mediation analysis frameworks, presenting the most up-to-date methodological insights that will drive the field of study forward. Thus, readers of this research article will have extensive background knowledge, likely already knowing about 'mediation analysis'.

Comparative register research has shown repeatedly that situational characteristics such as audience and purpose are associated with specific patterns of linguistic variation. Because linguistic features are functional, they are used to differing extents in different registers, conforming to the situational characteristics of those registers. For example, the pronoun *I* is extremely common in conversation because speakers talk a lot about their own actions and feelings. The pronoun *you* is common in conversation because there is a specific interlocutor who is physically present (see Biber et al. 1999: 333–334). The pronoun *you* is also relatively common in university course syllabi, because these texts are addressed to a specific readership (see Biber 2006a: 50–52). In contrast, neither pronoun is common in academic prose, because of the informational (rather than personal) communicative purpose, and the absence of a specific individual addressee.

Similar functionally-motivated linguistic patterns exist among the sub-registers of academic writing (see Biber 2006a: chapter 4; Gray 2015b). These

patterns correspond to differences in purpose and audience associated with specific registers (e.g., textbooks versus research articles), as well as differences across disciplines. For example, both past tense verbs and present tense verbs are frequent in humanities and education university textbooks; textbooks in these disciplines use past tense to narrate past events and provide historical background, while present tense is used for more informational characterizations and the discussion of ideas. Most verb phrases occur in the active voice, as in the following passage from a literary criticism textbook:

Text Sample 3.1
Literary criticism textbook

Barthes (see also chapter 4, pp. 70–71) was undoubtedly the most entertaining, witty and daring of the French theorists of the 1960s and 1970s. His career took several turns, but preserved a central theme: the conventionality of all forms of representation. He defines literature (in an early essay) as ‘a message of the signification of things and not their meaning (by “signification” I refer to the process which produces the meaning and not this meaning itself)’. He echoes Roman Jakobson’s definition of the ‘poetic’ as the ‘set to the message’, but Barthes stresses the process of signification, which appears less and less predictable as his work proceeds.

Raman Selden and Peter Widdowson. 1993.
‘A Reader’s Guide to Contemporary Literary Theory’
The University Press of Kentucky.

In contrast, engineering textbooks are focused mostly on processes and methodological discussions involving concrete entities. As a result, we see a much higher use of present tense verbs and a heavy reliance on passive voice verbs (promoting the object being manipulated, and demoting the agent doing the action), as in the following engineering textbook excerpt. For example:

Text Sample 3.2
Engineering textbook

Figure 7.4 shows how an involute curve can be generated. A string is wrapped around the base circle (a cylinder). As the string is unwrapped from the surface, a point on the string (point P) traces an involute profile. Figure 7.1 shows several gear types. For applications with parallel shafts, straight spur, stepped, helical, double helical, or herringbone gears are usually used. In the case of intersecting shafts, straight bevel, spiral bevel, or face gears are employed.

Arthur Erdman and George Sandor. 1990.
‘Mechanism Design: Analysis and Synthesis, Volume 1’
Prentice Hall.

When taking a register perspective to language variation, multiple levels of interpretation are possible, both in terms of defining registers and in describing the use of an individual linguistic feature. For example, at a more general level, textbooks and research articles are both realizations of 'academic prose'. Likewise, texts within the disciplines of humanities and engineering broadly represent 'academic prose'. Linguistically, passive voice verbs are important for distinguishing academic prose (whether textbooks or research articles, engineering or literary criticism) from conversation and popular written registers like fiction. That is, most written academic texts (regardless of specific sub-register or discipline) will use passive voice verbs to a greater extent than conversational or fictional registers.

At the same time, passive voice verbs turn out to be important for distinguishing among specific sub-registers and academic disciplines within academic writing: passives are especially common in engineering (and natural science) texts, but less common in humanities texts. Thus, just as we can identify sub-registers within the general 'academic writing' register, so too can we identify systematic patterns of linguistic variation amongst those sub-registers.

In the present chapter, we provide a linguistic description of modern academic writing at several different levels, beginning with a comparison of this general register to other general spoken and written registers in English (e.g., conversation, fiction, newspaper reportage), and then moving on to descriptions of more specific sub-registers of academic writing (e.g., textbooks versus research articles, and written texts from various academic disciplines). Our goal in the chapter is to provide a relatively comprehensive linguistic description of academic writing in the late twentieth century. That description serves as a background to the following chapters, which document the historical path taken to arrive at the present-day discourse style, along with the unique grammatical functions that have evolved in academic writing during that historical period.

3.2 Previous linguistic research on academic writing

Over the past 30 years, we have witnessed an explosion of research on academic discourse. This research has mostly been carried out by applied linguists interested in the linguistic characteristics of academic prose that differentiate it from other spoken and written registers (see, e.g., the extensive survey of research in Grabe and Kaplan 1996; and the chapters in the collections edited by J. Flowerdew 2002; Hewings 2001; Markkanen

and Schroder 1997; see also the survey of research in L. Flowerdew 2002). Although research articles have been the most widely studied register, other written academic registers that have been analyzed include textbooks (Biber 2006a; Biber, Conrad, and Cortes 2004; Conrad 1996; Freddi 2005; Moore 2002), PhD or master's theses/dissertations (Bunton 2005; Charles 2006a, 2006b; Samraj 2008), peer review reports (Fortanet 2008), 'comment' articles (Lewin 2005), book reviews (Groom 2005), and business reports (Yeung 2007).¹

Research on academic writing has investigated discourse characteristics on all linguistic levels, including the use of core grammatical features, the discourse functions of phraseological sequences, the rhetorical organization of academic texts, and complex patterns of linguistic co-occurrence. Our primary goal in this book is to provide detailed analyses of the distinctive grammatical patterns that distinguish academic writing from other registers. However, our review of previous research would be remiss to not mention two areas which have received considerable attention: vocabulary/phraseology and discourse structure. Thus, we briefly discuss these before moving on to research on the grammatical characteristics of academic writing.

Academic vocabulary has received considerable attention related to the development of word lists for language learners. These lists are compiled from analyses of corpora of academic texts, to identify the most important words used in academic writing (see, e.g., Coxhead 2000; Gardner and Davies 2014; see Huckin, Haynes, and Coady 1995; Nation 1990, 2001; Schmitt 2000; Schmitt and McCarthy 1997). Phraseological patterns have also garnered a great deal of attention. Some of these studies describe the collocations of particular words in academic texts (e.g., Williams 1998; Gledhill 2000; Marco 2000; Oakey 2002), while others have focused on longer recurrent sequences of three or more words (often referred to as 'lexical bundles'; see Hyland 2008; Biber et al. 1999: chapter 13; Biber, Conrad, and Cortes 2004; Cortes 2004; also see Ellis, Simpson-Vlach, and Maynard 2008, and Simpson-Vlach and Ellis 2010 on 'formulas').

Other academic writing researchers have taken a rhetorical or social/historical perspective to describe the discourse structure of written academic texts, and how the conventions of academic genres are shaped by the practices of researchers in particular discourse communities. Most of

¹ Student writing has also garnered a good deal of attention: Hardy and Römer 2013; Römer and Swales 2010; Nesi and Gardner 2012; Gardner and Nesi 2012; Grant and Ginther 2000; Green, Christopher and Mei 2000; Hinkel 2003; Jarvis, Grant, Bikowski, and Ferris 2003; Schleppergrell 1996; Spycher 2007; Altenberg and Granger 2001; Archer 2008; Flowerdew 2006; Loudermilk 2007.

these studies have focused on scientific or medical prose (see, e.g., the book-length studies by Atkinson 1999; Bazerman 1988; Berkenkotter and Huckin 1995; Gilbert and Mulkay 1984; Halliday and Martin 1993; Swales 1990; Valle 1999). Swales (1990) developed a framework for the analysis of ‘moves’ in academic research articles, and that framework has been widely applied to describe the structure of written academic genres across disciplines (e.g., see Koutsantoni 2006; Basturkmen 2009; Bhatia 1997; Brett 1994; Bruce 2008, 2009; Bunton 2005; Holmes 1997; Kanoksilapatham 2005; Lim 2006; Ozturk 2007; Samraj 2002, 2004, 2005; Stoller and Robinson 2013).

As this brief summary shows, considering the lexis/phrasology and discourse structure of written academic texts has been the primary focus of much previous research. The present book takes a complementary perspective, focusing on the distinctive grammatical features typical of academic writing. Previous studies on the grammatical characteristics of academic writing can be grouped into two major categories:

- (1) research focusing on the detailed analysis of a particular grammatical feature as it functions in academic writing
- (2) studies focusing on a range of grammatical features, either contrasting general academic writing with other broadly defined registers, or describing the linguistic patterns of variation among academic sub-registers (e.g., research articles from different disciplines)

In the following sub-sections, we review previous research falling into each of these categories.

3.2.1 *Studies of particular grammatical features in academic writing*

Previous linguistic analyses of academic writing have considered a wide range of individual grammatical features. Many of these are specialized features used for information packaging, such as:

- shell nouns and associated grammatical structures (Aktas and Cortes 2008; Gray and Cortes 2011; Gray 2010; Charles 2003)
- passive voice (Tarone, Dwyer, Gillette and Icke 1998; Baratta 2009)
- existential *there* (Huckin, Pesante, and Hutz 1988)
- extraposed constructions with anticipatory *it* (Groom 2005; Hewings and Hewings 2001)
- conditional clauses (Warchal 2010; Ferguson 2001)
- citations (Hewings, Lillis, and Vladimirov 2010; Hyland 1999; Charles 2006a,b)

Surprisingly, some conversational grammatical features are also employed in academic writing for information packaging functions. For example, Swales et al. (1998) and Hyland (2002a) describe the use of imperatives/commands in academic writing to focus the reader's attention on points of special importance, and Webber (1994) similarly describes the special functions of questions in written academic texts.

One area of research that has been especially prevalent in discussions of academic writing is the lexical and grammatical marking of stance and evaluation (e.g., Baratta 2009; Biber 2006a, 2006b; Biber and Finegan 1988, 1989b; Charles 2003, 2006b, 2007; Crompton 1997; Dressen 2003; Grabe and Kaplan, 1997; Holmes, 1986; Hunston 1993, 1994; Hyland 1994, 1996a,b 1998a,b; Hyland and Tse 2005; Mauranen and Bondi 2003; Meyer 1997; Myers 1989, 1990; Silver 2003; Salager-Meyer 1994; Swales and Burke 2003; Tucker 2003; Varttala 2003; see also Biber et al. 1999, especially chapter 12). Corpus-based approaches to stance typically focus on overt lexical and grammatical patterns that mark tentativeness, possibility, certainty, and attitudes in discourse. Stance investigations have analyzed language at a range of linguistic levels, including analyses of a single lexical item (e.g., Diani 2008 on *really* as an emphasizer, Aijmer 2009 on *seem*), lexical classes (e.g., Swales and Burke 2003 on evaluative adjectives), and particular grammatical structures (e.g., Biber and Finegan 1988 on stance adverbials; Baratta 2009 on passives; Charles 2006b, 2007, and Hyland and Tse 2005 on *that*-clauses; Hewings and Hewings 2002 on extraposed clauses).

Most research on stance in academic writing has focused on lexico-grammatical patterns; that is, on patterns in which a lexical item conveys a particular stance meaning relative to a proposition in the immediate discourse context. Stance studies rarely focus on evaluative lexis alone. Several of these studies attempt to document the entire system of stance marking, accounting for the range of stance meanings (attitudinal vs. epistemic meanings; certainty vs. doubt) as well as the range of grammatical devices used to express those meanings (e.g., modal verbs, adverbials, complement clauses). The frameworks presented in Biber (2006a, 2006b; Biber and Finegan 1989b; Biber et al. 1999) and Hyland (1996a, 1998a, 1998b) generally take this approach: they analyze a whole set of structural types of stance markers that convey a range of stance meanings.

Hyland (1996a,b, 1998b) has focused specifically on the expression of 'hedging' in academic research articles. For example, content-oriented hedges have two major functions: indicating the accuracy of a proposition (e.g., adverbials like *generally*, *approximately*, *partially*, *possibly*), or limiting

the writer's commitment to a proposition (e.g., *the present work indicates...*, *the model implies...*). Hyland (2002c) extends this line of research by investigating the ways in which authors refer to themselves (and when they do not refer to themselves) in academic prose. One of the themes that Hyland has developed over recent years is that academic research articles are interactive, in that authors actively try to involve the reader in the communication process. Specific studies in this line of research have investigated the use of addressee features (Hyland 2001), questions (Hyland 2002b), and directives (Hyland 2002a) in academic research articles.

A single grammatical structure that has received considerable attention in academic writing research is the noun phrase. As early as 1960, Wells documented the 'nominal style' of academic writing, in contrast to the 'verbal style' of other varieties. The most studied characteristic of the noun phrase in academic writing is the heavy reliance on nominalizations, described as the most important manifestation of 'grammatical metaphor' in modern science writing (see, e.g., Halliday 1988/2004; Halliday and Martin 1993; Banks 2005, 2008). Grammatical metaphor refers to the meaning expressed by nominalized forms, whereby the description of processes and actions – usually expressed with verbs – is converted into more static representations expressed by nouns (see Halliday 2004 for a collection of key works on nominalization and grammatical metaphor in science writing).

This focus on the nominalizations as a major building block of written academic texts has been extended in other studies to include associated noun phrase structures, such as nouns and adjectives as noun pre-modifiers, and prepositional phrases and participle clauses as noun post-modifiers (see Biber 1988; Biber et al. 1999; Biber and Clark 2002; Biber and Gray 2010; Biber, Gray, and Poonpon 2011; Fang, Schleppergrell, and Cox 2006; Vande Kopple 1994). In Section 3.3, we present some of the major corpus findings from previous research regarding the types of noun phrase structures that are particularly prevalent in academic writing, and then we return to more detailed analyses of many of these structures in the following chapters.

3.2.2 *A comprehensive survey of the distinctive grammatical features of academic writing: Contributions from The Longman Grammar of Spoken and Written English*

The research studies summarized in Section 3.2.1 focus on grammatical characteristics that have special functions in academic writing.

Surprisingly, though, some of these features are not actually common in that register. For example, previous studies have documented the special functions of questions, imperatives, stance adverbials, stance complement clauses, and conditional clauses in academic research writing, even though all of these grammatical features are used much more commonly in conversation than in writing. These features are salient in academic writing – maybe because they are more likely to be found in other registers – and they serve special discourse functions in academic writing. But in terms of their frequency of use, they are not typical of the grammatical style found in most written academic texts.

An alternative approach is to take a comparative corpus-based approach to identify the grammatical features that are used more commonly in academic writing than in other registers. In the present section, we summarize the findings from the most comprehensive study of this type: the corpus-based grammatical descriptions in the *Longman Grammar of Spoken and Written English* (LGSWE; Biber et al. 1999).

The LGSWE is a corpus-based reference grammar of English, describing the complete range of grammatical features in English, and comparing the distributions and functions of those features across four major registers: conversation, fiction, newspapers, and academic prose. The register comparisons are based on analysis of a representative corpus of texts (the Longman Spoken and Written English Corpus) containing approximately 5 million words from each register (see LGSWE: 24–35). The LGSWE describes the structural characteristics of grammatical features in English, but at the same time, it describes the patterns of use for those features: register differences and other contextual factors that influence the patterns of variation. For this reason, the LGSWE provides relatively comprehensive descriptions of those four registers. The description of academic prose is especially relevant for our purposes here.

The academic prose sub-corpus used for the LGSWE consists of both academic books (2.65 million words) and academic research articles (2.68 million words; see LGSWE: 32–34). Texts from both books and research articles were collected from a number of academic disciplines, including biology, chemistry, medicine, sociology, education, and law/history/politics. The corpus thus represents academic prose as a general register.

The frequency findings in the LGSWE allow us to identify core grammatical features that are typical of academic writing by virtue of the fact that they occur more frequently in that register. For example, corpus analysis in the LGSWE shows that nouns are especially common in academic

prose: there are about 300,000 nouns per million words in academic prose, compared to only around 150,000 per million words in conversation (Biber et al. 1999: 235). Adjectives are also much more common in academic prose than other registers (about 80,000 adjectives per million words in academic prose versus 20,000 in conversation and 60,000 in fiction and newspapers; Biber et al. 1999: 506). These grammatical features are not necessarily salient, and as a result, many of these patterns have not been the focus of previous investigations. Taken together, though, they provide a comprehensive description of the grammatical style typically employed in written academic discourse.

Table 3.1 catalogs the distinctive grammatical features of academic writing, based on a survey of all quantitative findings reported in the *LGSWE*. Specifically, Table 3.1 lists all grammatical features described in the *LGSWE* that occur much more frequently in academic prose than in the other registers (conversation, fiction, and newspaper prose).

The findings summarized from the *LGSWE* show that written ‘academic’ grammatical features come from most structural categories. However, three word classes are especially prevalent: nouns, adjectives, and prepositions. Overall, these grammatical classes are more frequent in academic prose than in other registers, and there are many related specific features that are especially characteristic of academic prose (e.g., nominalizations, noun phrases with multiple modifiers, stance noun + *of*-phrase). In contrast, verbs are generally much less common in academic prose than in other registers, although there are specific verb categories that are typical of academic prose (e.g., copula *be*, existence verbs, derived verbs, and passive voice verbs). Similarly, there are specific categories of adverbs and adverbials (e.g., linking adverbials) that are especially common in academic prose, even though adverbs overall are more common in spoken registers.

At the same time, few of these linguistic features are uniquely characteristic of academic prose. Rather, most of these characteristics occur in other registers as well, although they are much more common in academic writing. Features which are markedly more common in, but not restricted to, a particular register can be considered ‘register features’ rather than ‘register markers’ (see Biber and Conrad 2009: 53–55). That is, these features are (1) pervasive – distributed throughout a text from the register, and (2) frequent – occurring more commonly in the target register than in most comparison registers’ (Biber and Conrad 2009: 53). Taken together, these features provide an overall portrait of the distinctive grammatical landscape of academic writing.

Table 3.1 *Grammatical features that are especially common in academic prose (based on a survey of the Longman Grammar of Spoken and Written English)*

Feature	Pattern of use
Nouns and noun phrases:	
Nouns: overall pp. 65	Approximately 60% of all content words in academic prose are nouns
Nouns vs. pronouns pp. 235–236	Nouns are much more common than pronouns in academic prose, especially in object positions
Absence of pronouns pp. 235–236	Pronouns are generally rare in academic prose
Specific pronouns: <i>this</i> and generic <i>one</i> pp. 349–350, 354–355	Much more common in academic prose; <i>this</i> is used for immediate textual reference; <i>one</i> is used for generic rather than specific reference
Plural nouns pp. 291–292	Much more common in writing than in conversation; Most common in academic prose
Nominalizations pp. 322–323	Much more common in academic prose, especially nouns formed with <i>-tion</i> and <i>-ity</i>
Anaphoric expressions pp. 237–238	Anaphoric reference is usually expressed with a determiner + noun (rather than a pronoun)
Definite article <i>the</i> pp. 267–269	Much more common in writing than in conversation; Most common in academic prose
Demonstrative determiners p. 270, 274–275	Most common in academic prose; especially <i>this</i> and <i>these</i>
Noun phrases with modifiers p. 578	60% of all noun phrases in academic prose have a modifier
Noun phrases with pre-modifiers p. 589, 597	Very common in academic prose (and newspapers)
Nouns as pre-modifiers p. 589–596	Very common in academic prose (and newspapers) (e.g., <u>government agencies</u>)
Noun phrases with post-modifiers p. 606–608	Very common in academic prose (and newspapers)
Noun phrases with multiple post-modifiers p. 640–644	Most common in academic prose (e.g., <u>the utilization of such devices for social purposes</u>)
Noun <i>and/or</i> noun binomial phrases pp. 1033–1034	Most common in academic prose (e.g., <i>size and shape</i>)
Adjectives and adjective phrases:	
Adjectives: overall p.65, 506	Adjectives are much more common in academic prose than in conversation or fiction
Attributive adjectives p.506, 589	Much more common in academic prose (e.g., <u>the basic logical content</u>)

Table 3.1 (*cont.*)

Feature	Pattern of use
Specific predicative adjectives pp. 440	Several predicative adjectives are notably more common in academic prose than in other registers: <i>different, important, difficult, possible, necessary, available, useful</i>
Derived adjectives pp. 531–533	Much more common in academic prose, especially adjectives formed with <i>-al</i>
Verbs and verb phrases:	
Copula <i>be</i> Copular verb <i>become</i> pp. 359–360, 437–439	Most common in academic prose
“Existence” verbs pp. 366, 369, 419	Much more common in writing than in conversation; Most common in academic prose (e.g., <i>include, involve, indicate</i>)
Specific lexical verbs pp. 367–372	Several verbs are notably more common in academic prose than in other registers: Activity verbs: <i>use, produce, provide, apply, form, obtain, reduce</i> Communication verbs: <i>describe, suggest</i> Mental verbs: <i>consider, assume, determine</i> Causative / Occurrence / Existence verbs: <i>follow, allow, require, include, involve, contain, exist, indicate, represent</i>
Specific prepositional verbs (especially passive prepositional verbs) pp. 416–418	Several prepositional verbs are notably more common in academic prose than in other registers: Activity verbs: <i>deal with, BE applied to, BE used in, BE derived from</i> Communication verbs: <i>refer to</i> Mental verbs: <i>BE known as</i> Causative / Occurrence / Existence verbs: <i>lead to, result in, occur in, depend on, consist of, BE based on, BE associated with, BE related to</i>
Verbs with inanimate subjects pp. 378–380	Common only in academic prose (e.g., <i>such comparisons suggest. . .</i>)
Derived verbs pp. 400–403	Most common in academic prose, especially verbs formed with <i>re-</i> and <i>-ize</i> (e.g., <i>reabsorb, cannibalize, itemize</i>)
Tense and aspect pp. 456–462	Academic prose relies primarily on simple aspect, present tense verb phrases
Passive voice pp. 476–480, 937–940	Much more common in academic prose, especially the ‘short’ passive (with no <i>by</i> -phrase)

Table 3.1 (cont.)

Feature	Pattern of use
Specific passive verbs pp. 478–480	Several verbs are especially common with passive voice in academic prose; for example: <i>BE + made, given, taken, used, found, seen, considered, shown</i>
Adverbs and adverbials:	
Specific adverbs pp. 560–563	Several adverbs are notably more common in academic prose than in conversation: <i>often, usually, significantly, more, relatively, especially, particularly, generally, indeed</i>
Specific amplifiers pp. 560–563	A few amplifiers are notably more common in academic prose than in conversation: <i>extremely, highly</i>
Specific degree adverbs pp. 566–569	A few degree adverbs are notably more common in academic prose than in conversation: <i>relatively, fairly, slightly</i>
Linking adverbials pp. 766, 880–882	Most common in academic prose; especially <i>however, thus, therefore, for example (e.g.)</i>
Purpose and concessive adverbials pp. 784, 786, 820–821, 824–825, 826	Most common in academic prose (e.g., <i>in order to, although</i>)
Dependent clause features:	
Finite relative clauses with adverbial gaps pp. 624–625	Most common in academic prose; especially with the relativizer <i>in which</i> (e.g., <i>a mutant vimentin in which Ser82 is changed to S82E</i>)
Participle clauses as post-modifiers in noun phrases p. 606, 630–632	Very common in academic prose (and newspapers) (e.g., <i>the assumptions given above</i>)
Noun complement clauses with a <i>that</i> -clause p. 648–650	Most common in academic prose (e.g., <i>the fact that...; a possibility that...; no doubt that</i>)
Noun complement clauses with a <i>to</i> -clause p. 652–653	Very common in academic prose (and newspapers) (e.g., <i>an attempt to...; the ability to...</i>)
Abstract noun + <i>of</i> + <i>ing</i> -clause pp. 653–655	Most common in academic prose, especially with the head nouns <i>way, cost, means, method, possibility, effect, problem, process, risk</i> (e.g., <i>methods of assessing error</i>)
Extraposed <i>that</i> -clauses pp. 672–675	Most common in academic prose, especially controlled by the adjectives <i>clear, (un)likely, and (im)possible</i> (e.g., <i>it is unlikely that any insect exceeds this velocity</i>)

Table 3.1 (*cont.*)

Feature	Pattern of use
Extraposed <i>to</i> -clauses pp. 720–724	Most common in academic prose, especially controlled by adjectives (e.g., <i>(im)possible, difficult, hard, important, necessary</i>) (e.g., <i>It is important to specify the conditions. . .</i>)
Subject predicative <i>to</i> -clause pp. 714–715, 723	Common only in academic prose (and newspapers) (e.g., <i>The first step is to evaluate the expression</i>)
ing-clauses controlled by adjective predicates p.749	Most common in academic prose; (e.g., <i>capable of, important for/in, useful for/in: formalist strategies are useful for analyzing drama</i>)
Concessive adverbial clauses pp. 820–825	Most common in academic prose (and newspapers) (<i>though, although</i>)
Other features	
Prepositions p. 92	Most common in academic prose
<i>Of</i> -phrases pp. 301–302	Much more common in writing than in conversation; Most common in academic prose
Prepositional phrases as post-modifiers in noun phrases p. 606–608, 634–638	Very common in academic prose (and newspapers) (e.g., <i>the effect on the final state</i>)
Stance noun + <i>of</i> -phrase pp. 984–986	Most common in academic prose, especially <i>possibility of, value of, importance of, problem of, understanding of</i>
<i>that/those</i> + <i>of</i> -phrase pp. 307–308	Common only in academic prose
Preposition + <i>which</i> in relative clauses with adverbial gaps p. 624–626	Common only in academic prose, especially <i>in which</i> and <i>to which</i>
Selected coordination tags: pp. 116–117	Common only in academic prose, especially <i>etc.</i>
Quantifier <i>each</i>	Most common in academic prose
Semi-determiners <i>same, other, certain, and such</i> pp. 282–283	Much more common in academic prose
Dual gender reference: <i>he or she, his or her, he/she</i> pp. 316–317	Common only in academic prose
Lexical bundles with noun phrases and/or prepositional phrases pp. 997, 1015–1019	Very common in academic prose; e.g., <i>the end of the, the nature of the, one of the most, the way in which, the extent to which, the fact that the, as a result of, at the time of, in the case/absence/form/presence of, on the basis of, on the other hand</i>

3.2.3 Studies of register variation that include academic writing

Studies of register variation generally consider a wide range of grammatical characteristics, with the goal of comparing registers with respect to their exploitation of these features. Many of these studies compare the characteristics of academic writing with other spoken and/or written registers, while others describe the patterns of variation among sub-registers within the general category of academic writing.

For example, Biber (2006a) describes the typical linguistic characteristics of written academic registers in American universities (textbooks and course-packs) contrasted with spoken university registers (e.g., classroom teaching, advising sessions), and non-academic written registers found in the university (e.g., course syllabi, university catalogs). In addition to the physical mode difference (writing versus speech), these registers differ from one another with respect to interactiveness, production circumstances, communicative purpose, target audience, and so on. That study identifies a number of grammatical features that are considerably more common in academic written registers than in spoken registers or other written registers (e.g., common nouns, nominalizations, prepositional phrases, linking adverbials, passive voice, noun + *that*-clause constructions, and extraposed *to*-clause constructions).

Gray (2015b) takes this kind of research a step further by comparing and contrasting the grammatical characteristics of academic research writing from different disciplines, while simultaneously considering the influence of different types of research (i.e., qualitative, quantitative, and theoretical research). This study found that linguistic variation across six disciplines and three research types could be associated with a number of situational characteristics. For example, some of the patterns observed in the corpus could be explained by specific disciplinary differences (e.g., the frequent use of *if*-conditional clauses in philosophy; the frequent use of linking adverbials in theoretical physics). Other linguistic patterns were associated with the type of research. For example, regardless of discipline, qualitative research articles utilized narrative features (e.g., past tense verbs) of language to a greater extent than quantitative or theoretical research.

'Multi-dimensional' (MD) analyses – a research approach developed to investigate the linguistic patterns of register variation – have often included linguistic descriptions of academic writing compared to other spoken and written registers (see, e.g., Biber 1988, 1995, 2006a; Conrad and Biber 2001). Studies in this research tradition use large corpora of naturally-occurring texts to represent the range of spoken and written registers in a discourse domain. These registers are compared with respect to 'dimensions' of variation, comprising constellations of linguistic features that typically co-occur in texts.

MD analysis uses statistical factor analysis to reduce a large number of linguistic variables to a few basic parameters of linguistic variation. In MD analyses, the distribution of individual linguistic features is analyzed in a corpus of texts. Factor analysis is used to identify the systematic co-occurrence patterns among those linguistic features – the ‘dimensions’ – and then texts and registers are compared along each dimension. Each dimension comprises a group of linguistic features that usually co-occur in texts (e.g., nouns, attributive adjectives, prepositional phrases); these co-occurrence patterns are identified statistically using factor analysis. The co-occurrence patterns are then interpreted to assess their underlying situational, social, and cognitive functions.

Researchers have undertaken MD analyses of numerous languages and of many discourse domains in English. Given that each of these studies is based on a different corpus of texts, representing a different language and/or discourse domain, it is reasonable to expect that they would each identify a unique set of dimensions. However, despite these differences, there are certain striking similarities in the dimensions that are uncovered across these studies. Most importantly, in nearly all previous MD studies, there is a dimension associated with an oral-literate opposition (see Biber 2014). Table 3.2 summarizes the linguistic composition of this oral/literate dimension in studies of particular discourse domains in English.

Linguistically, this opposition is realized as two fundamentally different ways of constructing discourse: clausal versus phrasal. That is, the oral pole of this dimension consists of verb classes (e.g., mental verbs, communication verbs), grammatical characteristics of verb phrases (e.g., present tense, progressive aspect), and modifiers of verbs and clauses (e.g., adverbs and stance adverbials). In most studies, these ‘oral’ features also include various kinds of dependent clauses that function as clausal constituents, including adverbial clauses and finite complement clauses. These clausal features usually co-occur with pronoun classes and various types of colloquial features (like discourse markers). In contrast, the ‘literate’ pole usually consists of phrasal devices that mostly function as elements of noun phrases, especially nouns, nominalizations, attributive adjectives, and prepositional phrases.

Functionally, this ‘oral-literate’ dimension is usually interpreted as distinguishing between a personal/involved focus (personal stance, interactivity, and/or real time production features) versus informational focus. And in nearly every case, this parameter is the first dimension identified by the statistical factor analysis (i.e., it is the most important factor, accounting for the greatest amount of shared variance).

For our purposes here, the most important characteristic of this first dimension is that it consistently distinguishes academic writing (and other

Table 3.2 *The oral/literate dimension (Factor 1) in selected MD studies of particular discourse domains in English*

Discourse domain	Linguistic features defining Dimension 1	Register pattern along Dimension 1
General spoken and written registers; Biber (1985, 1986, 1988)	mental verbs, present tense verbs, pronouns: 1 st and 2 nd person, <i>it</i> , indefinite, demonstratives, emphatics, hedges, amplifiers, contractions, <i>that</i> -complement clauses, causative adverbial clauses, WH complement clauses <i>VERSUS</i> nouns, long words, prepositions, type/token ratio, attributive adjectives	face-to-face conversations, telephone conversations, personal letters <i>VERSUS</i> official documents, academic prose
World English spoken and written registers; Xiao (2009)	mental verbs, present tense verbs, adverbs, boosters, compromisers, pronouns: 1 st , 2 nd , 3 rd person, <i>it</i> , indefinite, demonstratives, contractions, <i>that</i> -deletion in complement clauses, causative adverbial clauses, WH nominal clauses <i>VERSUS</i> nouns, long words, nominalizations, prepositions, type/token ratio, attributive adjectives, passives	private speech, public speech <i>VERSUS</i> press reportage, instructional writing, academic writing
University spoken and written registers; Biber (2006a)	contractions, pronouns, mental / activity / communication verbs, present tense, progressive aspect, time / place / stance adverbials, WH-questions, <i>that</i> -clauses, WH-clauses, adverbial clauses <i>VERSUS</i> nouns, nominalizations, attributive adjectives, prepositional phrases, long words, passives, WH-relative clauses	service encounters, office hours, study groups, classroom teaching <i>VERSUS</i> textbooks, course packs, institutional writing
Elementary school spoken and written registers; Reppen (2001)	initial <i>and</i> , time adverbials, 3 rd person pronouns <i>VERSUS</i> nouns, long words, nominalizations, passives, attributive adjectives, prepositional phrases	student conversations, oral narratives <i>VERSUS</i> science and social science textbooks
ESL spoken and written exam responses; Biber and Gray (2013a)	mental verbs, present tense, modals, 3 rd person pronouns, <i>that</i> -clauses, adverbial clauses <i>VERSUS</i> nouns, attributive adjectives, prepositional phrases, long words, passives	spoken, independent tasks (low-scoring) <i>VERSUS</i> written, integrated tasks (high-scoring)

Table 3.2 (*cont.*)

Discourse domain	Linguistic features defining Dimension 1	Register pattern along Dimension 1
Eighteenth century speech-based and written registers; Biber (2001)	1 st and 2 nd person pronouns, present tense, possibility and prediction modals, <i>that</i> -deletion, mental verbs, emphatics <i>VERSUS</i> prepositions, passives, nouns, long words, past tense verbs	drama, letters <i>VERSUS</i> newspaper prose, academic prose, legal prose
Google text types; Biber and Kurjian (2007)	mental verbs, past tense, perfect aspect, communication verbs, pronouns: 1 st and 3 rd person, <i>that</i> -complement clauses, adverbial clauses <i>VERSUS</i> nouns	personal narrative <i>VERSUS</i> technical/informational discourse
Blogs; Grieve et al. (2011)	past tense, progressive verbs, activity verbs, emphatics, hedges, 1 st person pronouns, desire verb + <i>to</i> -clause <i>VERSUS</i> prepositions, attributive adjectives, nominalizations, passives, relative clauses	personal, narrative blogs <i>VERSUS</i> informational blogs
Academic research articles across disciplines; Gray (2013)	pronoun <i>it</i> , 1 st person pronouns, demonstrative pronouns, <i>be</i> and <i>have</i> as main verbs, causative verbs, modals of prediction, possibility, necessity; general adverbs, stance adverbials, adverbials of time; nouns of cognition, predicative adjectives, evaluative attributive adjectives; conditional adverbial clauses, <i>that</i> -clauses controlled by nouns and verbs of likelihood, <i>that</i> -clauses controlled by factive adjectives, <i>that</i> -clauses controlled by attitudinal and factive nouns, WH-clauses; <i>to</i> -clauses controlled by stance adjectives and verbs of probability <i>VERSUS</i> nouns, process nouns, past tense verbs, prepositions, type/token ratio, word length; passive post-nominal clauses, agentless passive voice verbs	theoretical philosophy <i>VERSUS</i> quantitative biology, quantitative physics

kinds of specialist informational writing) from all other registers. Further, within the domain of academic writing, this dimension distinguishes science research writing from other disciplines and other types of academic writing. Grammatically, these registers (i.e., academic writing generally, and science research writing in particular) are distinctive in their use of nouns and phrasal modifiers of noun phrases, in contrast to a greater use of clausal and involved grammatical features in other registers.

It is perhaps not surprising that Dimension 1 in the original 1988 MD analysis was strongly associated with the oral/literate opposition, given that the corpus in that study ranged from spoken conversational texts to written expository texts. For the same reason, it is somewhat predictable that a similar dimension would have emerged from the study of spoken and written registers in World English varieties (Xiao 2009), the study of eighteenth century general written and speech-based registers (Biber 2001), the study of university spoken and written registers (Biber 2006a), and the study of spoken and written responses on a standard English as a second language (ESL) exam (Biber and Gray 2013a).

However, it is more surprising that restricted comparisons of written registers would uncover a first dimension with a similar set of co-occurring linguistic features, associated with a similar opposition between oral-clausal and informational-phrasal styles (Reppen 2001; Gray 2013; Biber and Kurjian 2007; Grieve et al. 2011). In all of these cases, the linguistic composition of Dimension 1 is surprisingly similar, generally opposing verbs, dependent clauses, pronouns, and interpersonal features versus nouns and phrasal noun modifiers. Of course, the specific patterns of register variation vary, distinguishing between 'oral' versus 'literate' registers within a restricted discourse domain, and within a single mode. But the linguistic differences among those registers are strikingly similar across studies, suggesting that this is a fundamentally important parameter of register variation. In the following chapters, we undertake detailed linguistic investigations of many of the individual linguistic features from this dimension of variation.

3.3 Case studies illustrating the distinctive grammatical characteristics of academic research writing

3.3.1 Academic writing contrasted with conversation: Which is more grammatically complex?

Grammatical complexity is usually linked with elaboration and clausal embedding in linguistic theory. A 'simple' clause has only a subject, verb, and object or complement. A 'simple' noun phrase has a determiner and

head noun. Additions to these structures represent elaboration, resulting in 'complex' grammar. In particular, there is widespread agreement among grammarians that embedded clauses are an important type of grammatical complexity (often contrasted with 'simple' clauses; see, e.g., Huddleston 1984: 378; Willis 2003: 192; Purpura 2004: 91; Carter and McCarthy 2006: 489).

Academic writing is often claimed to be structurally complex in these terms. Thus, academic writing has been described as having longer sentences, longer 't-units' (a main clause plus all associated dependent clauses), 'longer and more complex clauses with embedded phrases and clauses' (Hughes 1996: 34), and a greater use of subordinate clauses (see, e.g., O'Donnell et al. 1967; O'Donnell 1974; Kroll 1977; Chafe 1982; Brown and Yule 1983).

In contrast, conversation has long been described as grammatically simple in these terms. Conversational participants share time and place, and they normally also share extensive personal background knowledge. As a result, pronouns and vague expressions are common, and referring expressions generally do not need to be elaborated in conversation. Because of these factors, conversational grammar is assumed to be generally not complex, employing 'simple and short clauses, with little elaborate embedding' (Hughes 1996: 33).

The following conversational excerpt illustrates many of these characteristics:

Text Sample 3.3 Conversation

Non-clausal utterances are marked in **bold**

- BARRY: I went to the Institute of Terror.
 WENDY: You went to where?
 BARRY: **The Institute of Terror.**
 ...
 WENDY: **Oh.**
 BARRY: It's pretty cool. You want to go? I've got free tickets.
 WENDY: Is it – it's a – how long is it going to be open?
 BARRY: **Until the thirty first.**
 WENDY: **Cool.** It's, it's an, it's actually pretty scary and stuff?
 BARRY: I wouldn't go so far as to say it's really scary.
 WENDY: But it's cool. [laugh]
 BARRY: **Yeah.** ... I'll go with you. I wouldn't pay for you or anything.
 WENDY: [laugh]
 BARRY: But I'll go with you.
 WENDY: It's expensive, isn't it? It's like five bucks.
 BARRY: Yeah, this one's six.
 WENDY: **The one down here?** And you have free tickets?
 BARRY: **Well, yeah.** ...
 WENDY: **Wow. Cool.**

This conversational excerpt illustrates the reliance on short, simple clauses, such as *it's pretty cool*, *I've got free tickets*, *I'll go with you*, and *it's like five bucks*. In addition, the excerpt illustrates many of the most salient characteristics of conversation, including hesitations, false starts, and short non-clausal utterances.

However, an equally important – but less often noticed – aspect of conversational grammar is that it is complex, relying heavily on dependent clauses. In fact, nearly any conversational interaction illustrates a dense use of dependent clauses. So, for example, [Text Sample 3.4](#) is taken from the same conversation as [Sample 3.3](#):

Text Sample 3.4
Conversation

Dependent clauses shown in **bold underline**

- WENDY: Why are you going there?
 BARRY: Oh, I have to do three different kinds of reviews in my senior seminar class and so, it's a good excuse **to go out for dinner**. <laughing.> I – you know – my teacher was suggesting **that we review a restaurant** so I was trying **to think of a good restaurant**.
 WENDY: I don't think **I would have gone there**.
 BARRY: Really? I've, I used to go there all the time **when I was ten** and I hadn't been there since then.
 WENDY: I don't remember **it being any good**.
 BARRY: Don't tell your mom that.
 WENDY: Well I don't remember ninety percent of **what I do**.
 BARRY: Why's that?
 WENDY: Big fuzzy black spots – alcoholism makes you **forget everything**.

Corpus evidence (see below) shows that [Sample 3.4](#) is typical of conversation generally. Thus, the stereotype of conversation as grammatically simple is misleading. In fact, dependent clauses are pervasive in conversation. For example, in [Sample 3.4](#), almost every conversational turn includes one or more dependent clauses.

It similarly turns out that our linguistic stereotypes about academic writing are only partly correct. On the one hand, it is possible to find some academic texts with a heavy reliance on embedded clauses, as in [Text Sample 3.5](#):

Text Sample 3.5

Academic writing: philosophy textbook

Embedded clauses marked with [____]

[**Even if propositional attitude accounts succeeded in their own terms**], they would not explain most of [**what should be explained**]

by a theory of emotion]. Propositional attitude theories are often presented [as if they were a simple consequence of the idea [that emotions involve the occurrence of mental states [which represent states of affairs in the world (states with “content”)]]].

...

[What is distinctive about the propositional attitude theory] is the interpretation [it gives to the words thought and belief]. The mainstream philosophical tradition [in which Lyons is located] assumes [that our everyday understanding of these notions is adequate for a theory of emotion].

Paul E. Griffiths. 2008.

‘What Emotions Really Are: The Problem of Psychological Categories’
University of Chicago Press.

However, many other academic texts employ surprisingly few embedded clauses. These texts are composed of long, complex sentences, but the complexity is associated with dependent phrases (embedded in noun phrases), rather than dependent clauses. Thus consider the following excerpt from a biology research article. While there are a few dependent clauses in this passage, the much more important characteristic is the pervasive use of phrasal modifiers:

Text Sample 3.6

Academic writing: biology research article

Noun-modifying prepositional phrases marked with UNDERLINED ITALIC CAPS; noun-noun sequences marked with *bold italics*

One explanation FOR the failure OF *invasion biology* AS a predictive science is the absence OF manipulative experiments IN the tradition OF modern *community ecology* (Kareiva 1996). ... *Biocontrol programmes* offer unparalleled opportunities to study the *invasion process* (Memmott et al. 1998) and the work presented here comprises a *6-year, large-scale, field experiment* based around a *weed biological programme* IN New Zealand.

A biological *control programme* FOR broom (*Cytisus scoparius* (L.) Link) began in New Zealand in 1981. Broom is a weedy legume that has been introduced deliberately to many countries and, like many *alien plants*, has become a serious weed OF *pasture, forest and conservation areas* (Williams 1981; Parsons & Cuthbertson 1992; Bossard & Rejmanek 1994). The *weed biological control agent*, *Arytainilla spartiophila* (Hemiptera: Psyllidae), was given high priority FOR introduction ...

J. Memmott, P.G. Craze, H.M. Harman, P. Syrett, and S.V. Fowler. 2004.

‘The effect of propagule size on the invasion of an alien insect’
Journal of Animal Ecology, 74(1): 50–62.

Detailed consideration of individual texts like [Samples 3.3–3.6](#) help us to notice discrepancies between the preconceptions of linguists regarding grammatical complexity versus the linguistic patterns found in actual spoken and written texts. However, it is difficult to determine the general patterns of use through such an approach; we have no way of knowing whether an individual text sample is representative of the entire register. In contrast, corpus-based analysis is ideally suited for this kind of research question (see, e.g., Biber, Conrad, and Reppen 1998; McEnery, Tono, and Xiao 2006). By basing analyses on large, representative collections of texts, it is possible to document the linguistic patterns that are actually representative of a register. And through such analyses, we often discover patterns of use that directly contradict our preconceived notions.

Research findings relating to grammatical complexity provide a classic case study of this type, showing how corpus-based research can uncover linguistic patterns of use that directly contradict our prior beliefs and expectations. The results presented below are adapted from Biber and Gray (2010) and Biber, Gray, and Poonpon (2011), based on analysis of large corpora of texts from academic research writing (the *20th Century Research Article Corpus*) and natural conversation (from the *Longman Spoken and Written Corpus*).

In [Chapter 2](#), we introduced the linguistic features in English that are related to grammatical complexity. As noted in [Section 2.5.2](#), complexity features can be compared along two grammatical parameters: *Parameter A – Structural Type*, and *Parameter B – Syntactic Function*. The corpus analyses presented below show that these two grammatical parameters are especially important for understanding the preferred distribution of complexity features in conversation versus academic writing.

The first parameter – structural type – refers to the ways in which information can be conveyed and elaborated through different types of grammatical structures: finite dependent clauses, non-finite dependent clauses, and dependent phrases. These structural types differ in the amount of grammatical information that they incorporate. Finite dependent clauses mark tense and usually include a grammatical subject, which often identifies the agent of the action described by the verb. Non-finite dependent clauses can mark aspect and voice (active or passive), but they do not mark tense or modality. In most cases, non-finite clauses also do not include a grammatical subject. Both finite and non-finite dependent clauses incorporate the full set of other clause elements, including objects and adverbials, specifying the semantic relations among those clause elements. Dependent phrases are at the opposite extreme from finite dependent clauses, including almost no overt grammatical

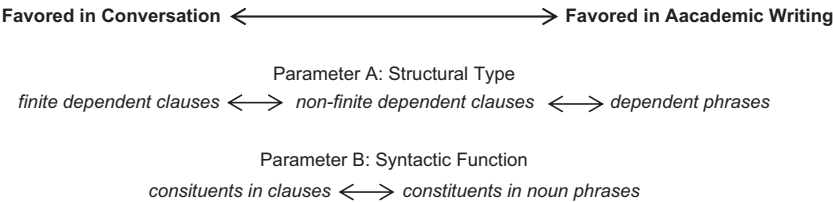


Figure 3.2 The distribution of complexity features in conversation and academic writing

information: no tense, aspect, voice, or modality; no overt identification of an agent; and no explicit identification of objects or adverbial elements.

The second parameter – syntactic function – refers to the syntactic role that the structure serves in a clause or phrase. The major distinction here is whether a structure functions syntactically as an element in a clause or as an element in a phrase. Clause-level elements can be adverbials (modifying the main verb) or complements (filling subject or object slots in a clause). Phrase-level elements modify the head of a phrase. These are most often modifiers of a noun (e.g., *hypothetical distribution*, *blood capillaries*, *the average for all groups*, *a decrease in opacity*), but they can also serve as modifiers of an adjective (e.g., *a relatively slow rate*).

The analyses in the following sections show that both of these grammatical parameters are important for understanding the distribution of features across spoken and written registers. Figure 3.2 (above) summarizes the patterns of variation in conversation versus academic writing: finite dependent clauses and constituents in clauses are more common in conversation, while dependent phrases and constituents within noun phrases are preferred in academic writing.

Tables 3.3–3.5 present the results of statistical comparisons for the use of specific complexity features in conversation versus academic writing. Table 3.3 presents the findings for finite dependent clause types; Table 3.4 displays non-finite dependent clause types; and Table 3.5 shows dependent phrase types. The rate of occurrence for each grammatical feature was measured in each text, allowing us to then calculate mean scores and standard deviations for each register (See Chapter 2, Section 2.4). Tables 3.3–3.5 report the results of statistical comparisons (using ANOVA) to determine whether the mean scores for conversation versus academic writing are significantly different from one another.

The r^2 scores in these tables provide a measure of the importance or strength of the difference between the two registers. For most of these

Table 3.3 *Statistical comparisons for finite dependent clause types*

Linguistic feature	Conversation mean score	Academic WR mean score	<i>F</i> value	Significance	<i>r</i> ²
Finite adverbial clauses					
total adverbial clauses	7.1	3.6	603.2	< .0001	.35
– <i>because</i> clause	2.0	0.6	336.9	< .0001	.23
– <i>if</i> clause	4.0	1.1	749.5	< .0001	.40
– <i>although</i> clause	0.6	0.05	777.6	< .0001	.41
Finite complement clauses					
verb + <i>that</i> -clause**	10.8	2.6	2196.7	< .0001	.66
verb + <i>WH</i> -clause	2.7	0.2	1413.9	< .0001	.55
adjective + <i>that</i> -clause	0.1	0.3	131.4	< .0001	.10
noun + <i>that</i> -clause	0.1	0.6	474.1	< .0001	.29
Finite noun modifier clauses					
<i>that</i> relative clauses	2.3	2.2	1.8	n.s.	
<i>WH</i> relative clauses	0.9	3.7	858.1	< .0001	.43

** including clauses with a ZERO complementizer

Table 3.4 *Statistical comparisons for non-finite dependent clause types*

Linguistic feature	Conversation mean score	Academic WR mean score	<i>F</i> value	Significance	<i>r</i> ²
Non-finite adverbial clauses					
<i>to</i> adverbial clauses	.08	.32	172.6	< .0001	.13
Non-finite complement clauses					
verb + <i>ing</i> -clause	1.3	0.2	842.5	< .0001	.42
verb + <i>to</i> -clause	4.7	3.4	166.6	< .0001	.13
adjective + <i>ing</i> -clause	0.04	0.1	48.6	< .0001	.04
adjective + <i>to</i> -clause	0.6	1.3	406.2	< .0001	.26
noun + <i>of</i> + <i>ing</i> -clause	0.05	0.4	310.6	< .0001	.21
noun + <i>to</i> -clause	0.9	2.8	856.8	< .0001	.43
Non-finite noun modifier clauses					
non-finite relative clauses	0.7	4.2	2257.3	< .0001	.66

features, the *r*² value is over .3 (i.e., over 30% of the variation in the feature can be predicted by the register difference between conversation and academic writing). However, the *r*² values for some features are considerably larger. For example, finite complement clauses controlled by verbs

Table 3.5 *Statistical comparisons for dependent phrase types (non-clausal)*

Linguistic feature	Conversation mean score	Academic WR mean score	<i>F</i> value	Significance	<i>r</i> ²
Adverbials					
Adverbs as adverbials	76.5	28.3	4581.6	< .0001	.80
Prepositional phrases as adverbials*	22.9	31.6	51.06	< .0001	.37
Noun modifiers					
Attributive adjectives	16.5	57.1	5787.8	< .0001	.84
Nouns as nominal pre-modifiers	19.0	57.4	1259.2	< .0001	.52
Total prepositional phrases as nominal post-modifiers*	6.3	51.9	1380.1	< .0001	.94
<i>of</i> -phrases as post-modifiers	4.6	34.1	9323.4	< .0001	.89
<i>in</i> -phrases as post-modifiers*	0.5	8.8	152.4	< .0001	.64
<i>on</i> -phrases as post-modifiers*	0.3	2.5	70.9	< .0001	.45
<i>with</i> -phrases as post-modifiers*	0.3	2.1	65.1	< .0001	.43
<i>for</i> -phrases as post-modifiers*	0.7	4.4	107.4	< .0001	.55

* based on a hand-coded sub-sample of 89 texts: 41 academic research articles; 48 conversations

(*that*-clauses and *WH*-clauses) have *r*² values over .55 (Table 3.3), while prepositional phrases as noun modifiers have an *r*² value around .90 (Table 3.5).

One especially interesting finding that emerges from these tables is that many of these complexity features are *not* common in academic writing. In fact, these linguistic differences are for the most part split between features that are strongly favored in conversation (and rare in writing) versus features that are strongly favored in academic writing (and rare in conversation). Figures 3.3 and 3.4 plot the most frequent features, showing the magnitude of these differences visually.

In terms of the two grammatical parameters introduced above, the most strongly favored types of structural complexity in conversation are finite dependent clauses (Parameter A) functioning as constituents in other clauses (Parameter B). Figure 3.3 plots the mean scores for finite clauses functioning as adverbials (*if* and *because*) and finite clauses functioning as verb complements (*that*, ZERO, and *WH*), showing that these clause types are much more common in conversation than in academic writing. Text Sample 3.7 illustrates the dense use of these finite clausal structures typical in everyday conversation.

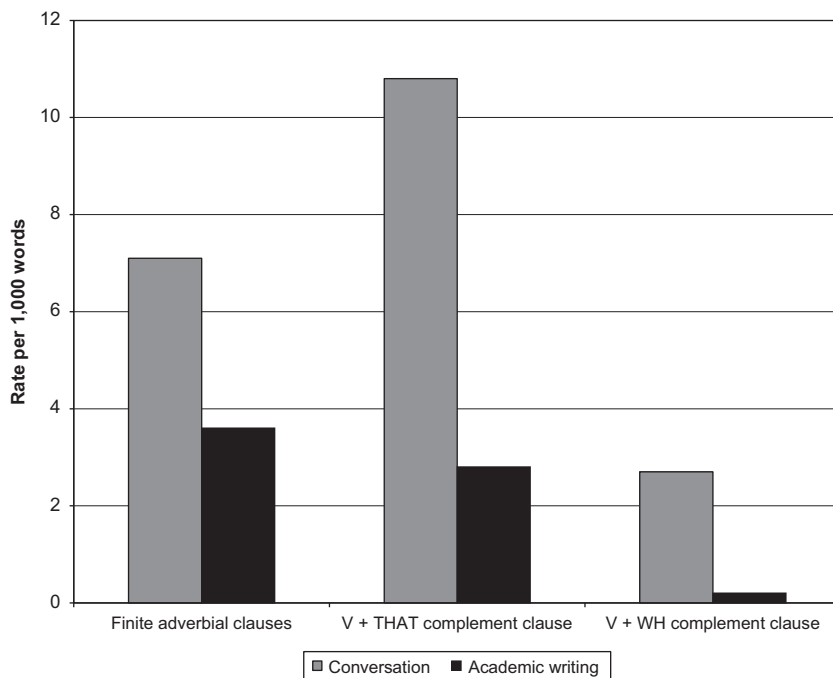


Figure 3.3 Common finite clause types functioning as clausal constituents

Text Sample 3.7

Conversation

Finite dependent clauses are **bold underlined**

GAYLE: And Dorothy said **Bob's getting terrible with, with the smoking.** Uh, he's really getting defiant about it **because there are so many restaurants where you can't smoke** and he just gets really mad and won't go to them.

...

PETER: Well they, they had a party. I forget **what it was.** They had it at a friend's house. I can't remember **why it wasn't at their house** any way. And they had bought a bottle of Bailey's **because they knew I liked Bailey's.**

...

GAYLE: I can't remember **who it was.** One of us kids.

...

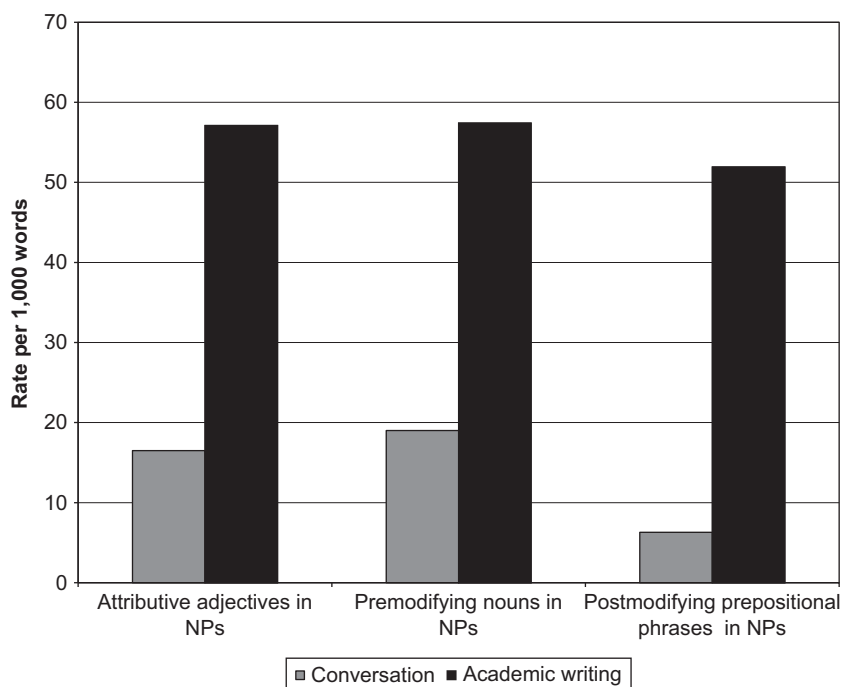


Figure 3.4 Common dependent phrasal types functioning as constituents in a noun phrase

- PETER: Oh. I'll tell you I think the biggest change in me is since I had my heart surgery.
- GAYLE: Really? Yeah I guess my, I mean I know my surgery was a good thing but
- PETER: <...> It makes you think. You realize it can happen to you.

In contrast, these two grammatical parameters are aligned in the opposite way in academic writig, which has a preferred use of phrasal rather than clausal structures (Parameter A) functioning as constituents in noun phrases (Parameter B). The use of prepositional phrases as post-nominal modifiers is the clearest case of this type (see Figure 3.2): they are extremely common in academic writing but rare in conversation. Attributive adjectives and nouns as nominal pre-modifiers are also phrasal constituents embedded in the noun phrase, and they pattern exactly like prepositional phrases as post-modifiers. Text Sample 3.8 illustrates the dense use of all three phrasal complexity features typical of academic prose:

Text Sample 3.8
Academic research article

Prepositional phrases functioning as noun modifiers are **bold underlined**; attributive adjectives are in *italics*; nouns as nominal pre-modifiers are in ***bold italics***

We expected that the use ***of different transformations*** would have *significant effects* **on our perceptions of spatial patterns in kelp holdfast assemblages**. Specifically, we were interested in the *qualitative ecological difference* **in emphasis between changes in composition vs. changes in relative abundance**. When analysing ***presence/absence*** data, the variability being measured is explicitly the variation **in the presence or absence of particular species (or taxa) in different holdfasts at different places**; thus *compositional change* is the *essential (and only)* feature. This can be contrasted against analyses based on other transformations (or *untransformed* data) for which variation **in relative abundance** plays a more *important*, or even a *dominant*, role **in the analysis**.

...

Similar patterns were seen for the analysis ***of assemblages*** based on ***Bray-Curtis*** dissimilarities ***of untransformed abundance data*** (Fig. 2b). Although the variation **from location to location** was detected as statistically significant for analyses **at the species, genus, family and order levels**, it was not statistically significant at the level **of class or phylum** using Bray-Curtis on *untransformed* data (Table 2). This result can be seen visually in ***NMDS*** plots, where *individual sites and locations* remain distinct for analyses **on untransformed data** using either species or families, but ***location-level*** differences are no greater than ***site-level*** differences **for the analysis of phyla** (see the ***left-hand*** side of Fig. 3a-c). For this and for *other* analyses using the ***Bray-Curtis*** measure, there was not, however, an *obvious* decrease **in the relative proportion of residual variation with decreases in taxonomic resolution** (Table 3).

Marti J. Anderson et al. 2005.

‘Relationships between taxonomic resolution and spatial
scales of multivariate variation.’

Journal of Animal Ecology, 74(4): 636–646.

Some grammatical complexity features are mixed or intermediate on the two parameters (see discussion in Section 2.5.2). That is, a feature may be ‘clausal’ in terms of structural type, but a ‘phrasal constituent’ in terms of syntactic function, and vice versa. These features include:

- a) finite clauses functioning as a constituent in a noun phrase
- b) phrases (non-clausal) functioning as a constituent in a clause
- c) non-finite clauses (all syntactic functions)

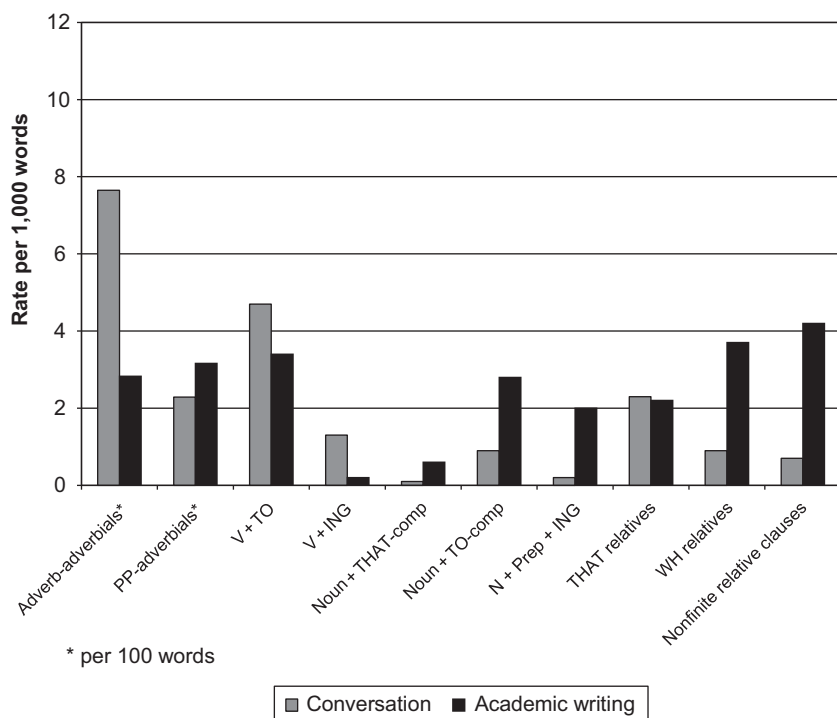


Figure 3.5 Dependent structures that mix the two parameters

Figure 3.5 shows that these intermediate features are generally less frequent in absolute terms than the other two sets of features, and that the differences between conversation and academic writing are generally less extreme. The exception is simple adverbs functioning as adverbials, which are very frequent in conversation. Other ‘mixed’ complexity features, however, show smaller differences between conversation and academic writing: prepositional phrases as adverbials are only moderately more frequent in academic writing than in conversation ($r^2 = .37$); verb + *to*-clause constructions are only moderately more frequent in conversation ($r^2 = .13$); *that* relative clauses occur with nearly the same frequency in both registers (no significant difference).

In addition, Figure 3.5 shows that Parameter B (syntactic function) continues to be particularly important for distinguishing between the characteristics of conversation and academic writing, even for dependent clauses. This is especially the case for dependent clauses functioning as constituents in a noun phrase, which tend to be strongly favored in

academic writing (except for *that* relative clauses), despite being clausal in terms of their structural type. Thus, noun complement clauses (*that* and *to*), *WH* relative clauses, and non-finite relative clauses are all much more common in academic writing than in conversation, although none of these structures is especially frequent in absolute terms. Text Sample 3.9 is a duplication of Text Sample 3.8, with these ‘mixed’ complexity features highlighted. This text sample is typical of academic writing generally, with seven dependent clauses functioning as noun phrase constituents versus only two functioning as clause constituents.

Text Sample 3.9
Academic research article

Dependent clauses as noun phrase constituents are **[bold underlined]**; dependent clauses as clause constituents are in *italics*.

We expected [*that the use of different transformations would have significant effects on our perceptions of spatial patterns in kelp holdfast assemblages*]. Specifically, we were interested in the qualitative ecological difference in emphasis between changes in composition vs. changes in relative abundance. When analysing presence/absence data, the variability **[being measured]** is explicitly the variation in the presence or absence of particular species (or taxa) in different holdfasts at different places: thus compositional change is the essential (and only) feature. This can be contrasted against analyses **[based on other transformations (or untransformed data)] [for which variation in relative abundance plays a more important, or even a dominant, role in the analysis]**.

Similar patterns were seen for the analysis of assemblages **[based on Bray-Curtis dissimilarities of untransformed abundance data]** (Fig. 2b). Although the variation from location to location was detected as statistically significant for analyses at the species, genus, family and order levels, it was not statistically significant at the level of class or phylum *using Bray-Curtis on untransformed data* (Table 2). This result can be seen visually in NMDS plots, **[where individual sites and locations remain distinct for analyses on untransformed data [using either species or families], but location-level differences are no greater than site-level differences for the analysis of phyla]** (see the left-hand side of Fig. 3a-c). For this and for other analyses **[using the Bray-Curtis measure]**, there was not, however, an obvious decrease in the relative proportion of residual variation with decreases in taxonomic resolution (Table 3).

Marti J. Anderson et al. 2005.

‘Relationships between taxonomic resolution and spatial
scales of multivariate variation’

Journal of Animal Ecology, 74(4): 636–646

Taken together, Figures 3.3 and 3.5 show that there are more dependent clauses in conversation than in academic writing. The difference is especially striking for finite dependent clauses functioning as clause constituents. For conversational participants, these kinds of structural elaboration do not seem complex; we generally do not even notice these structures, and they certainly do not inhibit normal communication. However, they are structurally ‘complex’ according to the definition of embedded clauses added on to simple clauses.

Our main goal in this section is not to claim that dependent clauses are unimportant in academic writing. As Figure 3.5 shows, certain types of dependent clause are actually more common in academic writing than conversation. However, the main point here is to draw attention to an often overlooked grammatical characteristic of academic writing: the dense use of phrasal modifiers. The discourse style of academic writing is dramatically different from conversation in its reliance on these phrasal modifiers. More surprisingly, we show in the following sections that academic writing is different from almost every other English register in the use of these features. This finding sets the stage for subsequent chapters, which trace the historical development of this distinctive grammatical style.

3.3.2 *University textbooks contrasted with university classroom teaching: Is there a general academic style?*

It is not surprising that academic writing should be different linguistically from conversation, given that the two differ situationally in just about every conceivable way (mode, production circumstances, communicative purposes, interactivity, etc.). However, the specific nature of the grammatical complexity differences between conversation and academic writing is certainly surprising, with dependent clauses being more frequent in conversation (and phrasal modifiers being much more frequent in writing).

Readers might be more surprised to learn that the clausal discourse style described in the last section is in fact the norm for most registers in English, characterizing all spoken registers as well as most written registers. In contrast, the phrasal discourse style introduced in the last section is restricted to written registers that have an informational purpose. In this section and the next, we present corpus findings that support these generalizations. First, we compare some of the typical grammatical

characteristics of written university textbooks to the characteristics of spoken university classroom teaching, showing that classroom teaching – despite its informational focus – is much more similar to conversation than academic writing. Then, in the next section, we compare the grammatical characteristics of academic writing to two other major written registers: fictional novels and newspaper prose.

We might predict that written university textbooks would be relatively similar to university classroom teaching in their characteristic grammatical features: both registers are addressed to the same audience (students), produced by the same types of speakers (academic professionals), and have similar communicative purposes (conveying information about academic topics). University classroom teaching differs from conversation in several respects: in a typical classroom lecture, the speaker and hearer have restricted interaction, the main speaker (the instructor) has the primary purpose of conveying information, and he/she has usually pre-planned the discourse. However, conversation and classroom teaching are similar in that the discourse in both situations is produced in real-time: the speaker is constructing utterances while she/he is actually speaking. This characteristic turns out to be important in influencing the grammatical discourse style of classroom teaching, despite the ways in which classroom teaching is otherwise similar to academic writing.

Biber (2006a) provides a detailed linguistic description of university textbooks and classroom teaching, compared to a range of other university registers (e.g., office hours, study group interactions, course syllabi, university catalogs). The results of that study show that textbooks and classroom teaching are sharply distinguished in the use of most lexico-grammatical features.

For example, Figure 3.6 shows the patterns of use for the four major content word classes, comparing classroom teaching and textbooks to conversation and academic research articles. Textbooks are not quite as extreme as research articles in their frequent use of nouns, but these two registers are nearly identical in their frequent use of adjectives, and their relatively infrequent use of verbs (when contrasted with the two spoken registers). Classroom teaching, on the other hand, is similar to conversation in its dense use of verbs, and its comparatively infrequent use of nouns.

Figure 3.7 plots the register distribution of three general types of finite dependent clauses: adverbial clauses (*if* and *because*), complement clauses

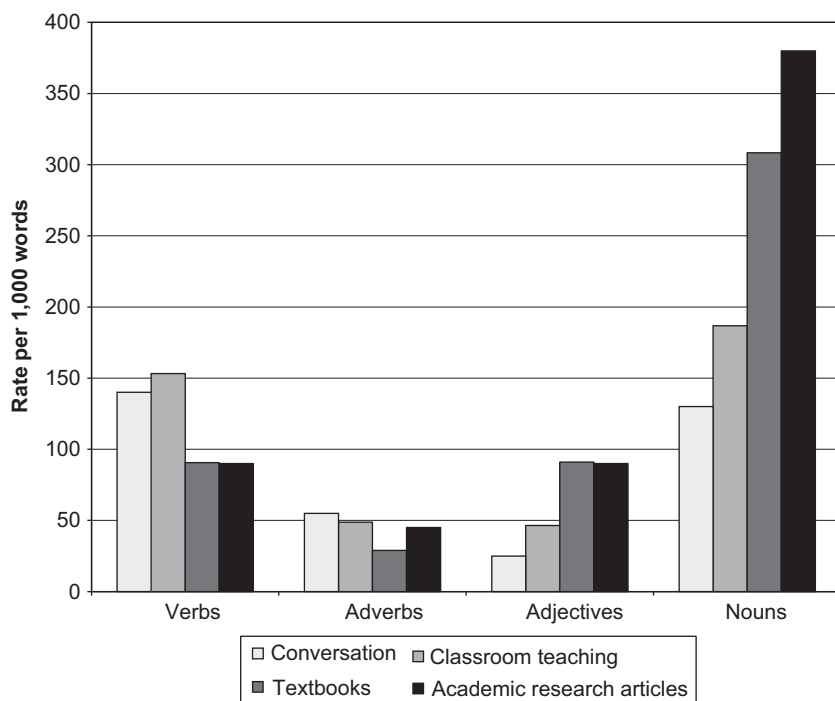


Figure 3.6 Major grammatical classes: academic speech and writing

(*that* and *WH*), and relative clauses (*that* and *WH*). Here again, classroom teaching is quite similar to conversation, with a dense use of adverbial clauses and complement clauses. Adverbial clauses are actually somewhat more common in classroom teaching, while complement clauses are more common in conversation. In contrast, both of these types of dependent clause are comparatively rare in academic writing (textbooks and research articles). Finite relative clauses show a different pattern, representing one of the few features that seems to conform to a general academic style addressed to students: very frequent in textbooks, but also frequent in classroom teaching. Interestingly, finite relative clauses are not especially frequent in academic research articles; we return to this finding in [Section 3.3.4](#).

[Text Sample 3.10](#) illustrates the typical discourse style of classroom teaching, with a heavy reliance on clausal structures, and relatively little use of phrasal modification:

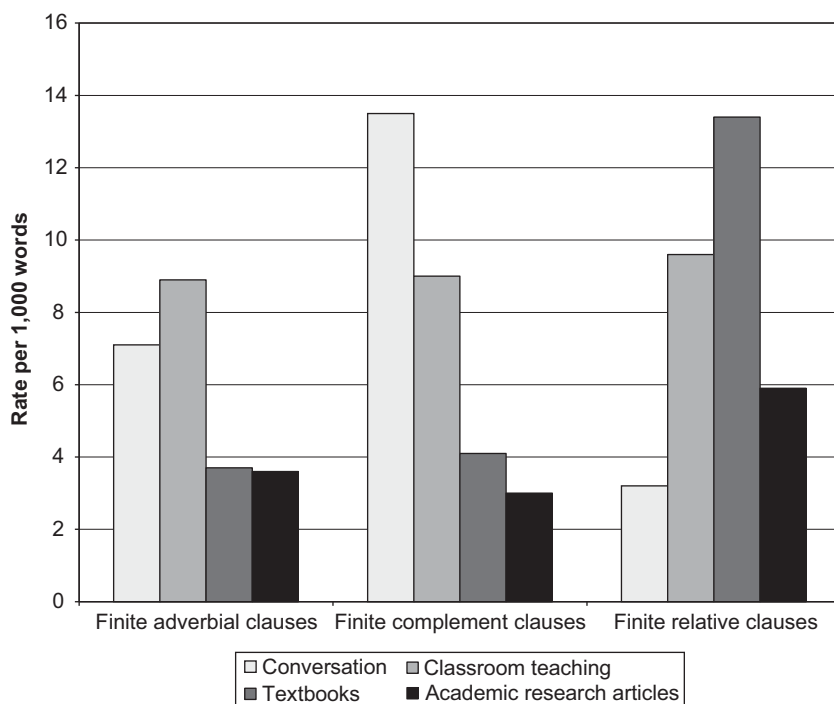


Figure 3.7 Finite dependent clauses: academic speech and writing

Text Sample 3.10

Classroom teaching; English

Pronouns are **bold underlined**; verbs are underlined italics; finite adverbial clauses and finite complement clauses are marked by [...].

Instructor: [What **I want you** to do in your free writes] is kind of reflect on [what do you think [**he means** here]]. Maybe – and [what **you could answer**] is would you want to live in that kind of place. Would you want to live there? And [if **you do**], Why? and do not, Why? And how does Rymmer give you clues? I think [Rymmer, especially in a poem like **this**, **he talks** about this hollowness at his core, sort of the absence of the bona fide, legitimate purpose to the whole thing]. I think [clues like **this** are embedded throughout that suggest [that Rymmer's pretty negative, or skeptical about this whole project]], right? And [what **I wanna know**] is, [if **you do want** to live there], why is that, and [if **you don't**], what is it about Rymmer's writing, or Rymmer's ideas that lead you to believe [that **you wouldn't want** to live there].

Both classroom teaching and textbooks are intermediate between the extremes of conversation and academic research articles. However, the grammatical style of classroom teaching (with its heavy reliance on clausal structures) is much more similar to conversation than academic writing. In contrast, textbooks are generally similar to academic research articles in their reliance on nouns and adjectives (and a lesser use of finite dependent clauses).

In sum, these findings show that there is no general academic style. Rather, there is a fundamental divide between speech and writing, with spoken academic texts being dramatically different from written academic texts in their reliance on a clausal rather than phrasal grammatical style. This section has compared two registers with a shared informational purpose but different modes. In the following section, we shift focus, restricting our analysis to only written registers but contrasting popular versus specialist registers.

3.3.3 *Grammatical features of academic writing contrasted with popular written registers*

The peculiar grammatical style of academic research writing becomes even more apparent when we compare it to the typical grammatical styles of other written registers. In the present section, we focus on two other general written registers: fictional novels and newspaper prose. Like academic prose, these registers are carefully written, revised, and edited, and they are addressed to large readerships. The primary differences among them relate to communicative purpose and the extent to which readers have specialized background knowledge:

	Fiction	Newspaper	Academic Writing
Purpose:	+entertainment	+informational	+informational
Audience:	+popular audience	+popular audience	+specialist audience

Fictional novels are written to entertain and edify, narrating a story, describing characters and places, portraying thoughts and emotions, and so on. Fiction is a complex written register, composed of prose sections interspersed with dialogue. It is not surprising that conversational dialogue in fiction is similar to natural conversation in its grammatical

characteristics. In the following, however, we show that the narrative and descriptive prose of fiction is also similar to spoken discourse, especially in its reliance on verbs, pronouns, and finite clauses.

Newspaper prose is more similar to academic writing in being informational in purpose. However, there are differences between the two: newspaper prose includes more narration and reported speech; academic prose includes more explanation and argumentation in support of its informational claims and descriptions. The two registers also differ in their audiences: newspaper prose is addressed to a large general audience. The author of a news story assumes that readers have some knowledge of the culture and current events, but beyond that, little background knowledge is taken for granted. In contrast, academic prose is written for specialists who share considerable expert background knowledge with the author.

Even the most basic grammatical features differ across these three registers, associated with the situational differences described above. For example, Table 3.6 presents the results of statistical comparisons for the four major word classes plus two specialized grammatical categories (passive voice verbs and nominalizations). Figure 3.8 plots the same information graphically. These three registers pattern in a highly regular manner with respect to these features: fiction makes the most frequent use of verbs and adverbs (being similar to conversation in this respect); academic prose makes the most frequent use of nouns, adjectives, and nominalizations; newspaper prose is consistently intermediate between the two extremes.

Table 3.6 *Statistical comparisons for major word classes across three general written registers*

Linguistic feature	ANOVA statistics	Mean Scores (per 1,000 words)		
		Novels	Newspaper prose	Academic prose
total verbs	$F = 487.5$; $p < .0001$; $r^2 = .69$	154.9	130.7	89.8
passive verbs	$F = 70.6$; $p < .0001$; $r^2 = .24$	9.9	12.2	18.2
adverbs	$F = 189.7$; $p < .0001$; $r^2 = .46$	64.2	41.9	28.8
common nouns	$F = 247.2$; $p < .0001$; $r^2 = .55$	220.5	314.5	381.7
nominalizations	$F = 139.5$; $p < .0001$; $r^2 = .40$	19.9	48.4	67.6
total adjectives	$F = 69.9$; $p < .0001$; $r^2 = .24$	56.6	75.3	90.0

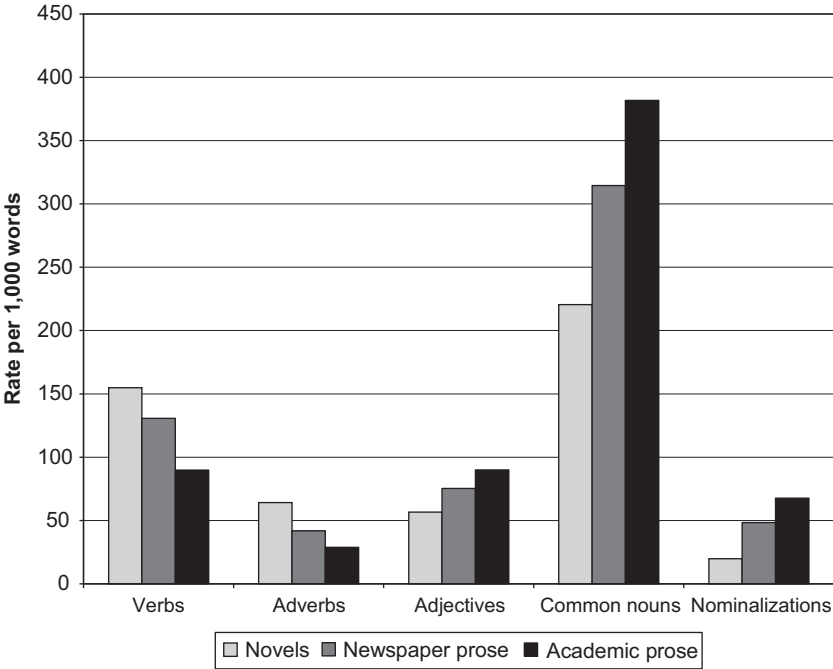


Figure 3.8 Major grammatical classes across three major written registers

Table 3.7 presents statistical results for the use of several specific dependent clause types, and Figure 3.9 plots the distribution of the major types. For the most part, we see the same pattern here: finite dependent clauses are most frequent in fiction, following again the pattern of use shown in Section 3.3.1 for conversation; finite dependent clauses are considerably less frequent in academic prose; newspaper prose is intermediate between the two extremes. The exception here is for the use of finite relative clauses, which are much more frequent in newspaper prose than in the other two registers. Overall, non-finite relative clauses are much less frequent than the other clause types, but they are the one type of dependent clause found most often in academic prose.

Fictional novels are surprisingly similar to conversation in the use of these grammatical features, despite the fact that they are monologic and produced in writing. It turns out that the clausal discourse style of fiction is typical of narrative prose as well as fictional dialogue. Thus consider the high density of verbs and dependent clauses in the following narrative text sample:

Table 3.7 Statistical comparisons for dependent clause features across three general written registers

Linguistic feature	ANOVA statistics	Mean Scores (per 1,000 words)		
		Novels	Newspaper prose	Academic prose
Finite adverbial clauses	$F = 35.4; p < .0001; r^2 = .15$	5.3	4.9	2.9
Complement clauses				
verb + <i>that</i> -clause	$F = 45.0; p < .0001; r^2 = .17$	6.7	5.9	3.3
verb + <i>WH</i> -clause	$F = 34.8; p < .0001; r^2 = .14$	1.4	0.4	0.2
noun + <i>that</i> -clause	$F = 3.7; p < .05; r^2 = .02$	0.3	0.6	0.5
verb + <i>to</i> -clause	$F = 57.5; p < .0001; r^2 = .21$	4.3	4.2	2.1
noun + <i>to</i> -clause	$F = 21.9; p < .0001; r^2 = .10$	0.5	1.2	0.4
Relative clauses				
<i>that</i> relative clauses	$F = 24.0; p < .0001; r^2 = .10$	2.3	4.3	2.8
<i>WH</i> relative clauses	$F = 64.7; p < .0001; r^2 = .24$	3.6	6.0	2.8
<i>-ed</i> (passive) relative clauses	$F = 22.5; p < .05; r^2 = .10$	1.9	2.3	3.5
<i>-ing</i> relative clauses	$F = 82.0; p < .0001; r^2 = .29$	0.2	0.3	1.0

Text Sample 3.11

Fictional prose

Verbs are in *bold italics*; dependent clauses are [underlined in brackets]

Penelope **was** glad [that Claire **had left** Frizzle], [for he **made being** alone not exactly alone]. The dog **loved** her and **offered** the nearest thing to human companionship. At least he **was waiting** [when she **returned** home from work in the evenings]. Aside from Frizzle [to **help** [**combat** the chilling silence of the apartment]], she **brought** home papers [to **grade**] and **plunged** into [**completing** her book on John Milton].

We have claimed in the preceding sections that the most distinctive grammatical characteristic of academic prose is its reliance on phrasal noun modifiers (contrasting with the rare use of clausal features). Table 3.8 and Figure 3.10 further support this generalization in comparison to the discourse styles of fiction and newspaper prose. Attributive adjectives are more common in fiction than in conversation (compare Figure 3.4), but they are only half as frequent as in academic prose. Nouns as pre-modifiers are rare in fiction but extremely common in academic prose. And the same pattern holds for *of*-phrases and other prepositional phrases as post-nominal modifiers. Newspaper reportage is again intermediate in the use

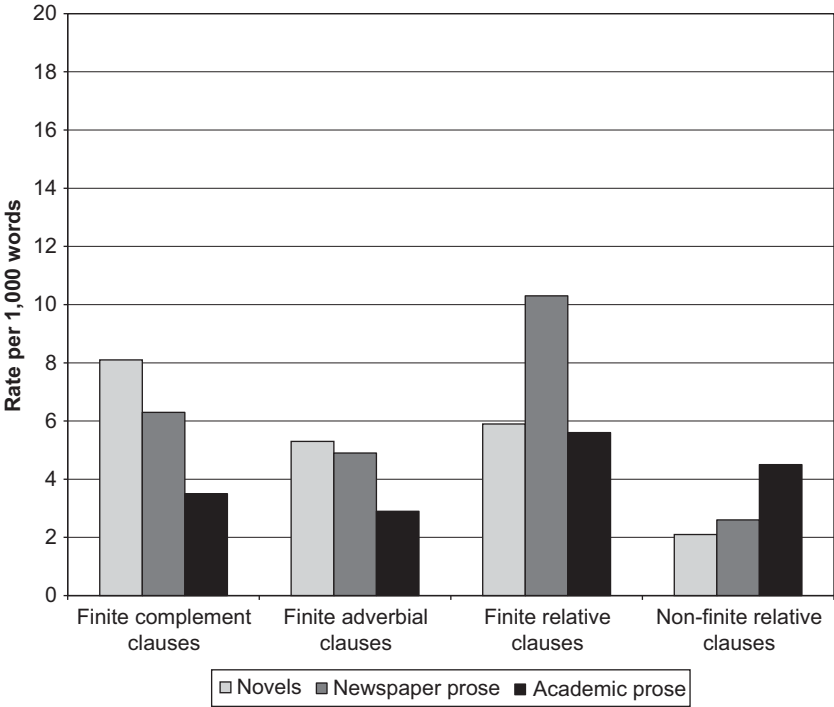


Figure 3.9 Dependent clause types across three major written registers

of these features, using phrasal modifiers more than many other registers of English, but not nearly to the same extent as in academic prose.

Newspaper prose combines narrative and informational communicative purposes. As a result, it employs many of the same grammatical devices as academic writing, although it uses those devices less commonly. For example, noun-noun sequences are prevalent in newspaper prose in comparison to written registers like fiction, as in [Text Sample 3.12](#):

Text Sample 3.12
Newspaper report

Noun-noun sequences are underlined; prepositional phrases as noun modifiers are in *italics*]

The Prime Minister declared her opposition [*to avenging the Locerbie air disaster*] yesterday, as American intelligence chiefs admitted failing to link the crash with any terrorist organization.

...
 With the incoming Bush administration [*in America*] certain to face pressure [*for retaliation*], Mrs Thatcher's outspoken rejection [*of reprisal raids*] could pose the first difficulties [*in her relationship [with the new President]*].
 ...

Meanwhile, a further tightening [*of baggage inspection procedures*] is likely to emerge from a review [*of Britain's airline and airport security*]. It will be launched this week at a meeting [*of the national Aviation Security Committee*], comprising government, airline, union and safety officials.

Philip Webster, Harvey Elliott, and Christopher Thomas. 1989.
 'Thatcher says no to revenge for Lockerbie'
The Times, 63, section N9, 1–18.

Prepositional phrases as noun modifiers are also prevalent in many news reports, although again not as frequent as in academic prose texts.

In summary, we have illustrated in this section how communicative purpose and audience have a major influence on grammatical discourse style within the written mode. Thus, popular written registers with

Table 3.8 *Statistical comparisons for phrasal modifiers across three general written registers*

Linguistic feature	ANOVA statistics	Mean Scores (per 1,000 words)		
		Novels	Newspaper prose	Academic prose
prepositional phrases as adverbials*	$F = 7.8$; $p < .001$; $r^2 = .15$	29.3	35.4	26.9
attributive adjectives	$F = 102.9$; $p < .0001$; $r^2 = .32$	30.5	49.3	65.2
nouns as nominal pre-modifiers	$F = 271.6$; $p < .0001$; $r^2 = .57$	14.3	35.2	67.3
of-phrases as noun modifiers	$F = 206.2$; $p < .0001$; $r^2 = .50$	14.4	19.1	31.5
other prepositional phrases as noun modifiers	$F = 16.0$; $p < .0001$; $r^2 = .27$	7.4	11.0	18.4

* These findings regarding the distribution of prepositional phrases are based on analysis of four prepositions – *in*, *on*, *for*, *with* – in ninety texts published in 1985. Ten instances of each preposition were randomly selected from each text for the analysis, to determine whether they were functioning as adverbials or noun modifiers. The rates of occurrence for each text were then extrapolated, based on the overall counts for those prepositions in the texts. See Biber and Gray (2012) for more details.

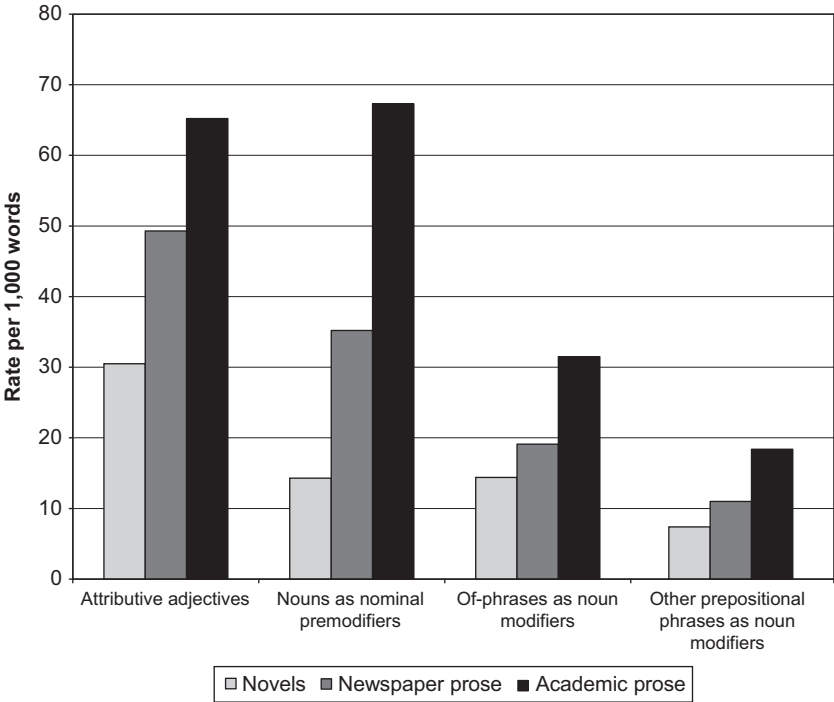


Figure 3.10 Noun phrase modifiers across three major written registers

non-informational purposes (e.g., fiction) use clausal linguistic styles, similar to spoken registers. Academic written registers, with informational purposes and specialist audiences, have developed a distinctive grammatical style, employing a dense use of nouns and phrasal modifiers rather than verbs and clauses. Newspaper prose is intermediate in this regard, using both clausal modifiers as well as phrasal modifiers, reflecting the fact that newspapers are written for a large, generalist readership, but for relatively informational purposes.

For our purposes in the present book, the main point of the present section has been to highlight the distinctive grammatical characteristics of academic prose. That is, even in comparison to an informational written register like newspaper prose, academic prose is distinctive in its infrequent use of clausal modifiers and its extremely frequent use of phrasal modifiers.

3.3.4 Grammatical variation among written sub-registers from different academic disciplines

Academic writing can be analyzed at many different levels of generality. For example, the descriptions in Section 3.2.2 (surveying corpus findings from the LGSWE) were based on analysis of a large general corpus of academic writing, including books and research articles from many different disciplines (see Biber et al. 1999: 32–34). At the other extreme, studies like Gray (2013, 2015a, 2015b) compare different disciplines, even distinguishing between sub-registers within those disciplines (i.e., qualitative, quantitative, and theoretical research articles).

In the present section, we illustrate the ways in which specific academic sub-registers vary in their use of grammatical features. To do so, we focus on differences among specialist research articles from three major discipline families: humanities, social science, and natural science (based on analysis of the 20th Century Research Article Corpus; see Chapter 2). In addition to disciplinary differences, we also consider the influence of audience, by comparing science research articles written for specialists to science writing published in multi-disciplinary venues addressed to a wider multi-disciplinary audience (i.e., the ‘Specialist Science’ sub-corpus of the 20th Century Research Article Corpus vs. recent articles published in the *Philosophical Transactions* and in *Science*; see Table 2.2 in Chapter 2).

Tables 3.9–3.11 present the results of statistical comparisons for specific grammatical features across these sub-registers of academic research

Table 3.9 *Statistical comparisons for major word classes across disciplines in academic writing*

Linguistic feature	ANOVA statistics	Mean Scores (per 1,000 words)			
		Humanities	Social science	Popular science	Specialist science
total verbs	$F = 46.2$; $p < .0001$; $r^2 = .40$	103.1	95.9	95.2	75.4
passive verbs	$F = 21.8$; $p < .0001$; $r^2 = .24$	13.3	19.7	19.2	18.7
adverbs	$F = 65.1$; $p < .0001$; $r^2 = .49$	37.7	28.5	33.1	22.1
common nouns	$F = 113.5$; $p < .0001$; $r^2 = .62$	314.3	372.8	340.9	423.0
nominalizations	$F = 7.0$; $p < .001$; $r^2 = .09$	62.8	70.5	72.1	61.0
total adjectives	$F = 5.0$; $p < .01$; $r^2 = .07$	90.5	87.0	98.9	87.0

Table 3.10 *Statistical comparisons for dependent clause features across disciplines in academic writing*

Linguistic feature	ANOVA statistics	Mean Scores (per 1,000 words)			
		Humanities	Social science	Popular science	Specialist science
Finite adverbial clauses	$F = 5.4; p < .001; r^2 = .07$	3.2	3.4	3.0	2.3
Finite complement clauses					
verb + <i>that</i> -clause	$F = 22.3; p < .0001; r^2 = .25$	4.0	3.2	5.0	2.2
adjective + <i>that</i> -clause	$F = 25.7; p < .0001; r^2 = .27$	0.5	0.2	0.9	0.1
noun + <i>that</i> -clause	$F = 14.0; p < .0001; r^2 = .17$	1.0	0.6	0.6	0.4
Finite noun modifier clauses					
<i>that</i> relative clauses	$F = 8.2; p < .0001; r^2 = .11$	3.6	3.1	3.3	2.1
WH relative clauses	$F = 34.5; p < .0001; r^2 = .34$	5.7	2.9	3.3	2.4
Non-finite complement clauses					
verb + <i>to</i> -clause	$F = 27.5; p < .0001; r^2 = .29$	3.2	2.4	2.4	1.3
adjective + <i>to</i> -clause	$F = 11.0; p < .0001; r^2 = .14$	1.0	0.7	1.1	0.5
noun + <i>to</i> -clause	$F = 11.5; p < .0001; r^2 = .14$	1.0	0.5	0.5	0.2
noun + <i>of</i> + <i>ing</i> -clause	$F = 11.4; p < .0001; r^2 = .14$	1.3	1.5	0.8	0.6
noun + Prep + <i>ing</i> -clause	$F = 5.9; p < .001; r^2 = .08$	0.7	1.1	1.0	0.7
Non-finite noun modifier clauses					
- <i>ed</i> (passive) relative clauses	$F = 3.3; p < .05; r^2 = .05$	2.9	3.6	3.5	3.3
- <i>ing</i> relative clauses	$F = 16.4; p < .0001; r^2 = .19$	0.4	0.9	1.1	1.0

writing. As the tables show, all of these grammatical features are associated with significant differences across the disciplines, and there are actually large differences in the use of some of these features (shown by the large r^2 scores).

Table 3.9 displays descriptive statistics for the major lexical word classes across disciplines, while Figure 3.11 presents the same information graphically. Verbs and adverbs are both important features here, shown by the r^2 scores greater than 0.4. Both of these features follow the same pattern: most common in humanities research writing; somewhat less common in popular science and the social sciences; and least frequent in specialist science research articles.

Common nouns have the opposite distribution, and the differences are even larger (reflected in the r^2 score of 0.62). In this case, the four registers follow a steady cline: least frequent in humanities writing; somewhat more frequent in popular science articles; more frequent again in social science articles; and by far the most frequent in specialist science writing.

As noted in Section 3.2.1, some researchers have focused on nominalizations as an especially important characteristic of academic writing (see Halliday 1988, 2004; Banks 2008). However, Figure 3.11 shows

Table 3.11 *Statistical comparisons for phrasal noun modifiers across disciplines in academic writing*

Linguistic feature	ANOVA statistics	Mean Scores (per 1,000 words)			
		Humanities	Popular science	Social science	Specialist science
attributive adjectives	$F = 5.1; p < .01; r^2 = .07$	65.5	72.9	61.1	63.9
nouns as nominal pre-modifiers	$F = 139.6; p < .0001; r^2 = .67$	24.3	56.7	66.4	76.6
prepositional phrases as adverbials*	$F = 13.1; p < .0001; r^2 = .37$	70.4	42.2	56.2	52.5
of-phrases as noun modifiers	$F = 16.7; p < .0001; r^2 = .20$	36.6	35.5	29.8	30.2
other prepositional phrases as noun modifiers*	$F = 8.3; p < .0001; r^2 = .27$	21.4	21.5	29.4	31.5
noun- <i>xxing</i> + noun	$F = 2.8; p < .05; r^2 = .05$	0.1	0.5	0.3	0.4
noun- <i>xxed</i> + noun	$F = 5.9; p < .001; r^2 = .10$	0.2	0.6	0.8	1.6

* These findings regarding the distribution of prepositional phrases are based on analysis of twenty texts from each sub-register. We sampled all prepositions from a 1,000-word segment of the research article, taken from the beginning, middle, and latter parts of the text. Then, each prepositional phrase was coded by hand to determine whether it was functioning as an adverbial or noun modifier. The rates of occurrence for each text were then extrapolated, based on the overall counts for those prepositions in the texts. Counts for prepositional phrase as adverbials include prepositional verb constructions as well as other adverbials.

that nominalizations are not especially frequent in absolute terms, and that there are only minor differences among academic sub-registers ($r^2 = .09$). Surprisingly, popular science employs more nominalizations than any other sub-register, while specialist science research articles make the least use of these devices. Similar to nominalizations, adjectives (including both attributive and predicative functions) are somewhat more common in popular science than in the other registers.

Table 3.10 focuses on the use of dependent clauses in these registers, presenting detailed information on several specific clause types. Figure 3.12 plots the distribution of five major dependent clause types. Although non-finite relative clauses are exceptional (see below), the other four dependent clause types are distributed in very similar ways: all four clause types are frequent in humanities writing, and most types are also frequent in popular science. These dependent clause types have intermediate

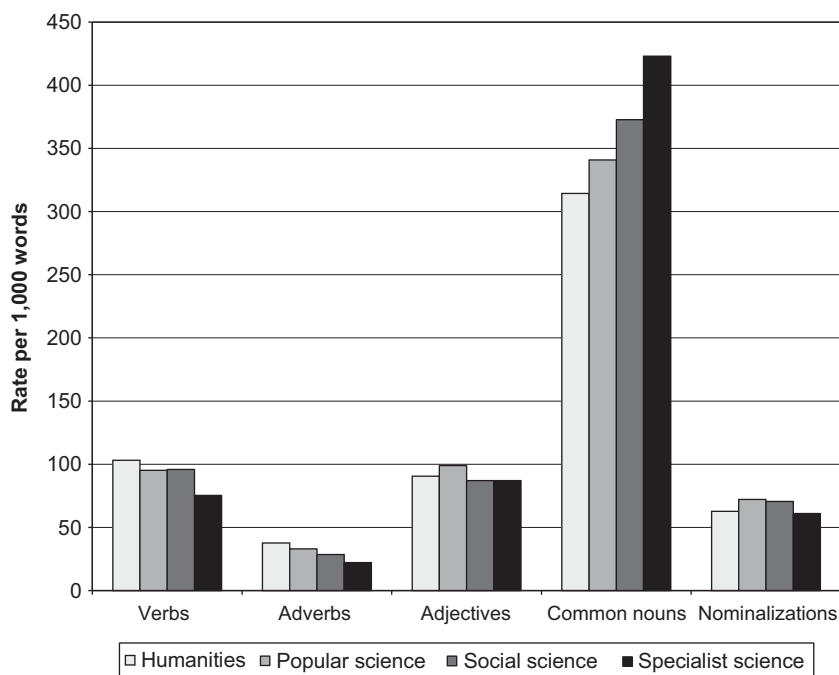


Figure 3.11 Major grammatical classes across sub-disciplines of academic writing

frequencies in social science. At the other extreme, though, all four dependent clause types are much less common in specialist science research articles.

Beyond those general patterns, each type of dependent clause has a slightly different distribution across sub-registers. Finite adverbial clauses are not very common overall in any of these registers, but they are even less common in specialist science. Complement clauses – both finite and non-finite – are frequent in both humanities and popular science writing. These clause types are less common in social science, but they are especially rare in specialist science. The majority of the complement clauses in academic prose are controlled by verbs, often used for reporting previous research claims (e.g., *Jones suggests that . . .*). However, Table 3.10 shows that complement clauses controlled by nouns follow the same distributional pattern: *that*-clauses controlled by nouns (e.g., *the fact that. . .*) and *to*-clauses controlled by nouns (e.g., *the attempt to. . .*) are most common in humanities; intermediate in popular science and social science; and rare in specialist science. Finite relative clauses are also comparatively rare in

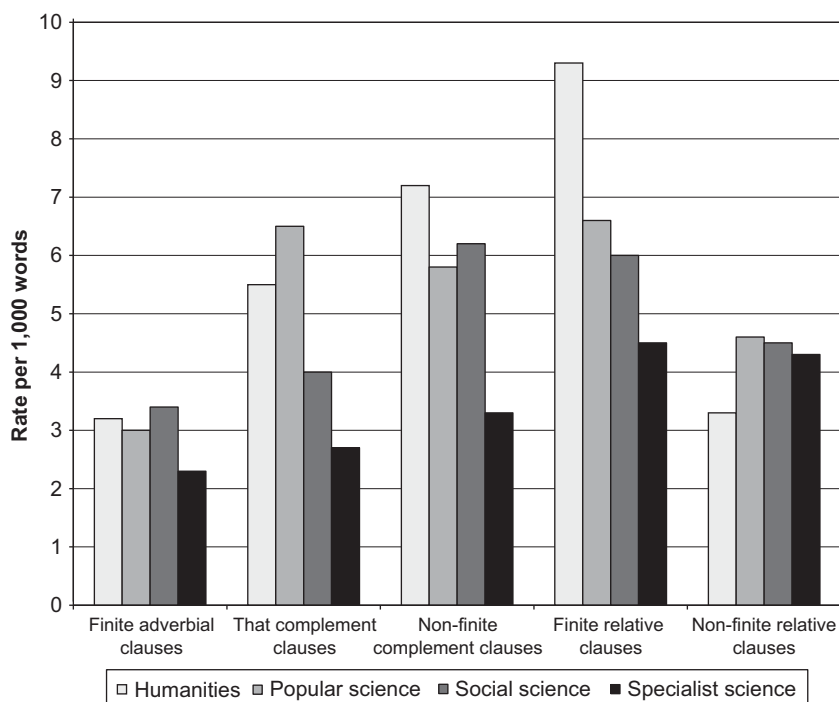


Figure 3.12 Dependent clause types across sub-disciplines of academic writing

specialist science, but relatively frequent in humanities writing. These differences are especially large for WH-relative clauses (see Table 3.10), but the same patterns hold for *that*-relative clauses. Both popular science and social science are intermediate in the use of those features.

As noted above, non-finite relative clauses are the only type of dependent clause to follow a different distributional pattern: nearly as frequent in specialist science as they are in social science and popular science, but notably less common in humanities writing. Linguistically, this dependent clause type is relatively similar to the phrasal modifiers favored in ‘literate’ discourse: it is minimally clausal (incorporating only a non-finite verb) and it functions syntactically as a noun modifier. It is thus noteworthy that this is the only dependent clause type to be strongly favored in specialist science writing.

Figure 3.13 presents the patterns of use for two features that function as pre-nominal modifiers: attributive adjectives and pre-nominal nouns (see also Table 3.11). Attributive adjectives are slightly more common in

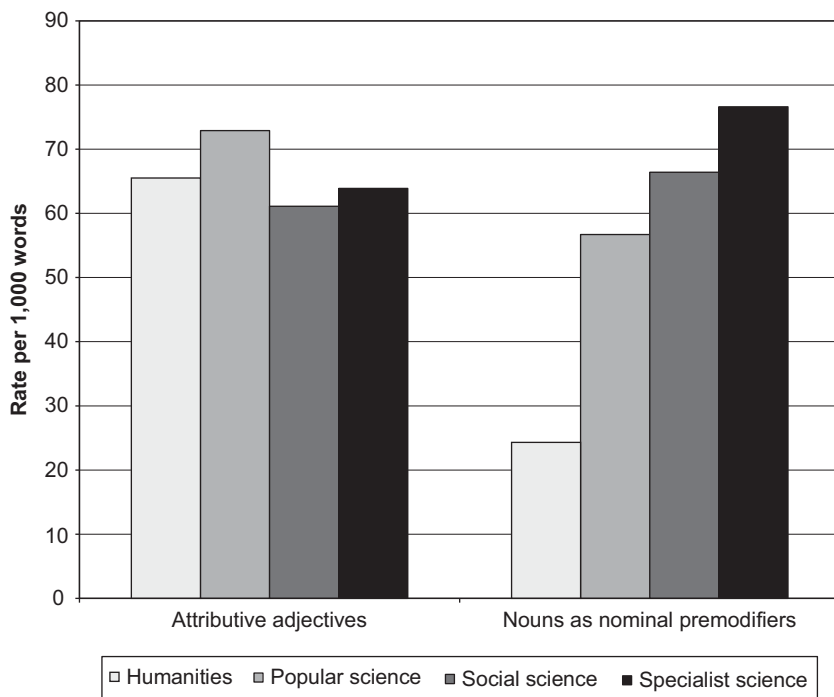


Figure 3.13 Noun phrase modifiers across sub-disciplines of academic writing

popular science articles than the other registers, but the differences are not statistically important ($r^2 = .07$). In contrast, there are major statistical differences ($r^2 = .67$) across disciplines in the use of nouns as nominal pre-modifiers: rare in the humanities; relatively common in popular science and social science; and most frequent in specialist science.

As Figure 3.14 shows, there are also major differences across academic sub-registers in the use of prepositional phrases (see also Table 3.11). Three grammatical types/functions are distinguished here: prepositional phrases functioning as adverbials (or as part of prepositional verbs), *of*-phrases functioning as noun modifiers, and other prepositional phrases functioning as noun modifiers. Prepositional phrases functioning as adverbials are considerably more common in humanities prose than in the other disciplines; they are least common in popular science writing. Prepositional phrases as noun modifiers are considerably less common. Humanities academic prose and popular science articles pattern alike in the use of these features: *of*-phrases functioning as noun modifiers are relatively common, while other

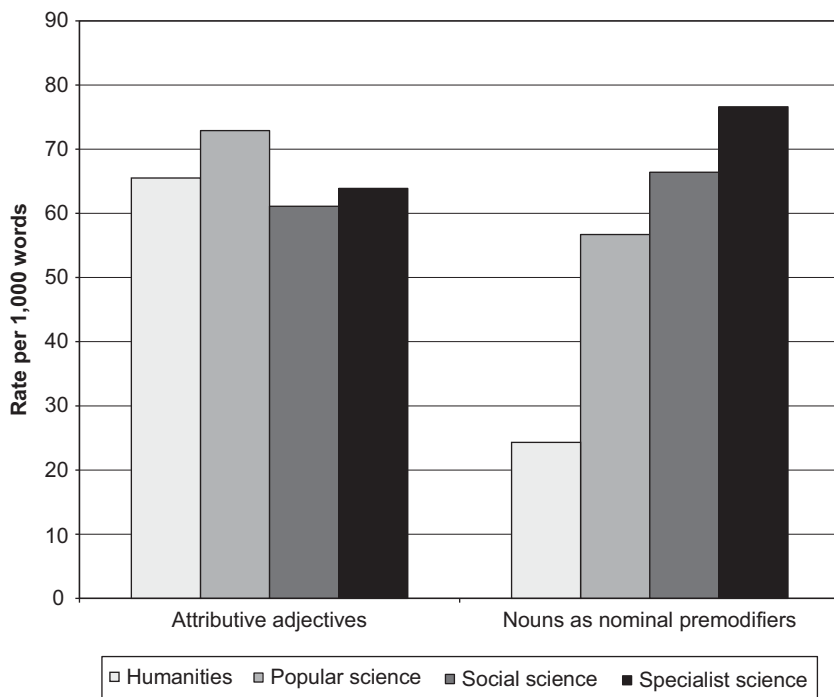


Figure 3.13 Noun phrase modifiers across sub-disciplines of academic writing

popular science articles than the other registers, but the differences are not statistically important ($r^2 = .07$). In contrast, there are major statistical differences ($r^2 = .67$) across disciplines in the use of nouns as nominal pre-modifiers: rare in the humanities; relatively common in popular science and social science; and most frequent in specialist science.

As Figure 3.14 shows, there are also major differences across academic sub-registers in the use of prepositional phrases (see also Table 3.11). Three grammatical types/functions are distinguished here: prepositional phrases functioning as adverbials (or as part of prepositional verbs), *of*-phrases functioning as noun modifiers, and other prepositional phrases functioning as noun modifiers. Prepositional phrases functioning as adverbials are considerably more common in humanities prose than in the other disciplines; they are least common in popular science writing. Prepositional phrases as noun modifiers are considerably less common. Humanities academic prose and popular science articles pattern alike in the use of these features: *of*-phrases functioning as noun modifiers are relatively common, while other

prepositional phrases functioning as noun modifiers are considerably less frequent. Text Sample 3.13 illustrates the dense use of prepositional phrases as adverbials or verb complements (shown in **bold underline**) and genitive *of*-phrases (shown in **BOLD CAPS**) in humanities research writing:

Text Sample 3.13
History research article

As a survey **OF** the relevant historical texts indicates, it is difficult to exaggerate the magnitude **OF** Zhou Dunyi's contribution. Zhou is often regarded **as** one **OF** the five Song masters who first formulated the Neo-Confucian "vista and determined its direction." Zhou Dunyi is appraised **as** "the pioneer" who "laid the pattern **OF** metaphysics and ethics for later Neo Confucianism." Wing-tsit Chan claims that Zhou Dunyi's originality consists **in** assimilating "the Taoist element **OF** non-being **to** Confucian thought," now carefully removed **from** "the fantasy and mysticism **OF** Taoism." If Neo-Confucianism is generally to be distinguished **from** earlier forms **OF** Confucian thought **by** its adaptation **OF** Daoist and Buddhist views, then Zhou Dunyi should be deemed **as** the first to demonstrate how fruitful this could be. **Through** Zhou Dunyi's efforts Confucian thought finally acquired an appropriately positive metaphysical foundation, one that bases Confucian moral teaching **on** an ontology that inspired successive generations **OF** Neo-Confucians' **in** their philosophical reflections.

Robin Wang, 2005.

'Zhou Dunyi's Diagram of the Supreme Ultimate Explained (Taijitu):
A Construction of the Confucian Metaphysics'
Journal of the History of Ideas, 66(3): 307–324.

Specialist science and social science research articles show a quite different pattern from humanities writing: *of*-phrases are considerably less common in science writing than in humanities, but other prepositional phrases functioning as noun modifiers are considerably more common in science articles than in humanities writing. Thus, *of*-phrases and other prepositional phrases functioning as noun modifiers occur with roughly the same frequency in science research writing (see Figure 3.14). Two prepositions are notably common in this role: *in* and *for*. Prepositional phrases headed by *in* are especially common in specialist science writing, while prepositional phrases headed by *for* are most common in specialist social science writing; for example:

Post-nominal modifiers with **IN**, specialist science research writing:

In addition, we analyzed skewing of *X inactivation* **IN** individual cells. ...

We generated multiple hamster rbip constructs with *mutations* **IN** this site. ...

The nucleolus is *a common stress sensor* of pivotal importance **IN** the p53 response.

We included in the spatial analysis only the environmental component of the residuals, correcting for *expected change* **IN** population size. ...

Post-nominal modifiers with **FOR**, specialist social science research writing:

The analysis sample **FOR** the current study consisted of 772 children.

One explanation **FOR** this relationship is that smaller classes are a key part of the capacity. ...

These activities ... could also represent *an opportunity* **FOR** fun together.

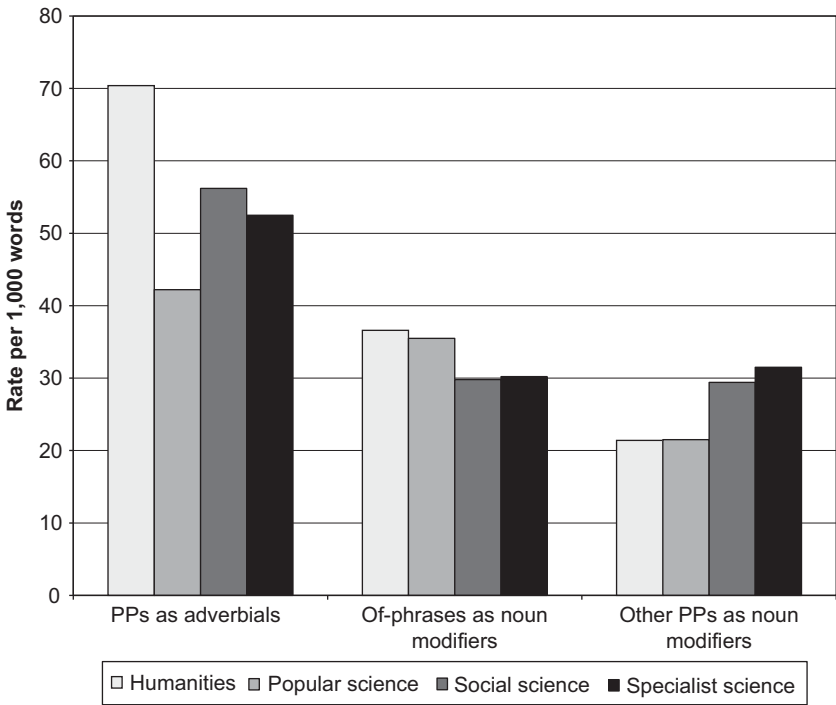


Figure 3.14 Prepositional phrases across sub-disciplines of academic writing

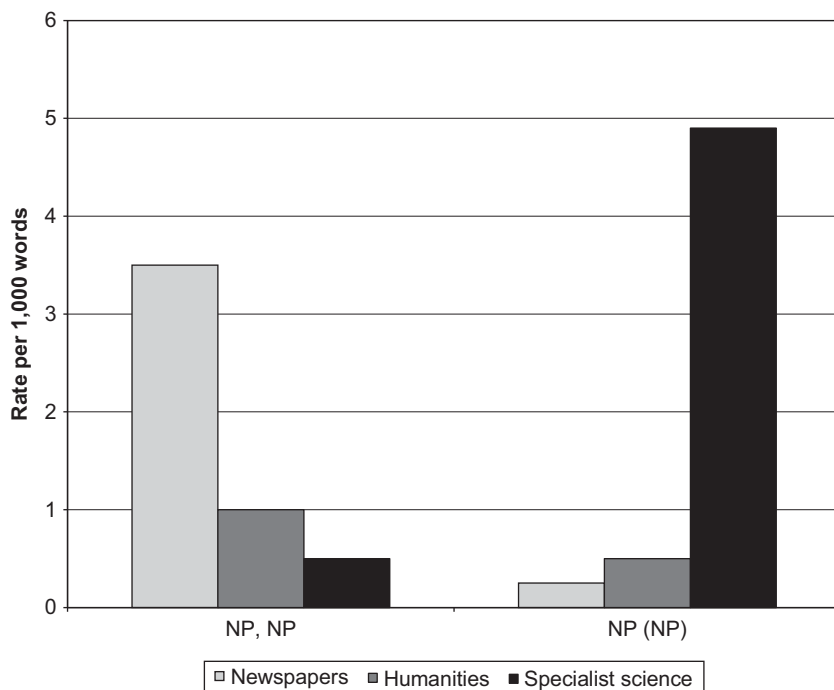


Figure 3.15 Appositive noun phrases in newspapers versus two sub-disciplines of academic writing

Appositive noun phrases are another phrasal device used for post-nominal modification. These structures are familiar to all readers of modern newspaper prose, usually expressed as two noun phrases separated by a comma: NP, NP, as in:

"A lot of companies are looking to get out of these plans anyway," says Dan McGinn, a pension consultant in Anaheim, Calif.

Figure 3.15 shows that NP, NP appositives are considerably more common in newspaper prose than in academic research writing. The most common functions of these structures are to describe the role or importance of an individual person, place, or institution; for example:

"You can buy anything and everything in the market," says Park, a trader who sells televisions she brings in from China.

All-terrain vehicles and biscuits were delivered to Banda Aceh, the capital of Aceh province.

The Indonesians also lashed out at Virginia-based Worldhelp, a missionary group founded by a Baptist disciple of Jerry Falwell.

While appositives with the format of *NP*, **NP** are not especially common in academic prose, *NP* (**NP**) appositives are very frequent. However, as Figure 3.15 shows, there is considerable variability among academic sub-registers in the use of appositives. Humanities research writing makes moderate use of *NP*, **NP** appositives, but to a considerably lesser extent than in newspaper reportage; *NP* (**NP**) appositives are relatively rare in humanities prose. In contrast, *NP* (**NP**) appositives are extremely frequent in science research writing. These structures can present many different types of information. For example:

We present the results of the International Meta-analysis of Mortality Impact of Systemic Sclerosis (**IMMISS**) ...

In four cohorts (**Athens, Keio, Mayo, and Florence**), investigators stated that. ...

Numerous variables were measured, including case status, sex, race, date of enrollment (**date of first visit to the cohort with the pertinent diagnosis**), age at first visit ...

We return to a more detailed discussion of the meanings and discourse functions of appositive noun phrases in Chapter 5.

Finally, Table 3.11 compares academic sub-registers for their use of a grammatical device that has been mostly disregarded in previous research: noun-participle compounds functioning as nominal pre-modifiers. These structures consist of either a pre-modifying noun combined with an *-ing* participle (e.g., *filament forming ability*) or a pre-modifying noun combined with an *-ed* participle (e.g., *Cdkl-induced vimentin-Ser55 phosphorylation*). The use of these structures enables an extremely dense packaging of information, as in:

This study considers the role of the cytoplasmic segments of Sed61a in the ribosome-binding and translocation-promoting activities of the Sec61 complex.

As Table 3.11 and Figure 3.16 show, these structures are not frequent in absolute numbers, but they are considerably more common in science research articles than in the other registers. In particular, constructions

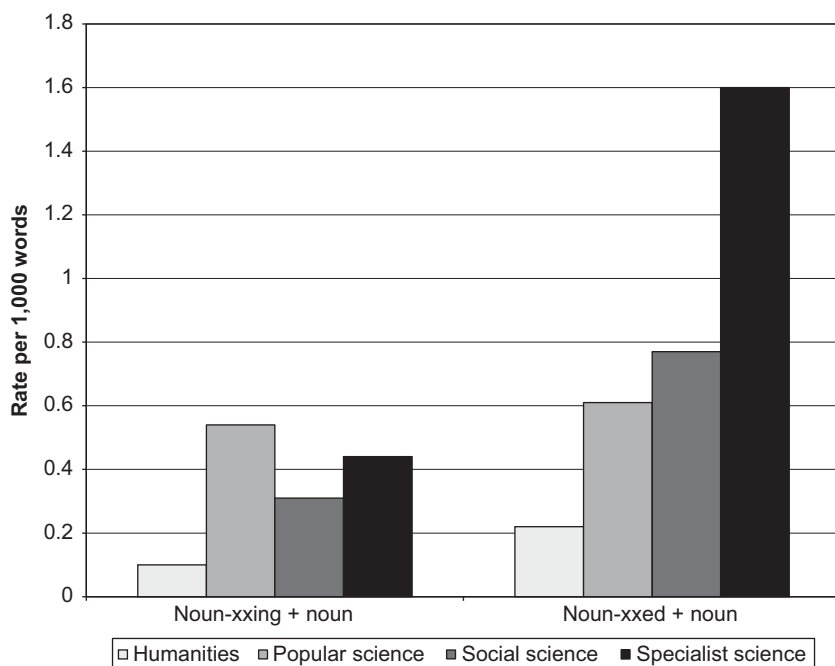


Figure 3.16 Noun + participle as nominal pre-modifiers across sub-disciplines of academic writing

with *-ed* participles are especially common in science research writing, as in:

Moreover, yeast centromeres cluster near the membrane-embedded spindle pole body.

Capan-1 cells were grown on coverslips, fixed with cold methanol, and stained with affinity-purified anti-centrobin.

Surprisingly, many of the grammatical features described in this chapter can be associated with a loss of explicitness. That is, as grammatical structures become more compressed, there is less overt expression of the meaning relationships among constituents.

One clear example of this pattern involves two alternative ways of indicating the logical relationship between clauses: linking adverbials (e.g., *however*, *therefore*) versus colons. Linking adverbials might be considered maximally explicit, in that they express the specific logical

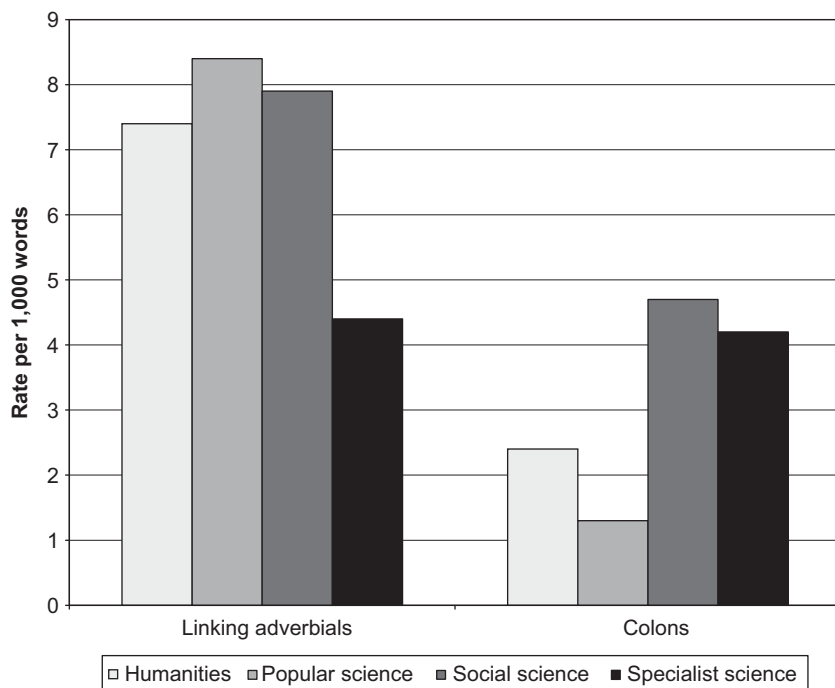


Figure 3.17 Explicit versus implicit signals of logical relations across sub-disciplines of academic writing

relationship between two clauses (e.g., contrast or consequence). In contrast, colons (and semicolons) are maximally inexplicit; they simply identify the existence of a logical relationship between two clauses, with no overt indication of what that relationship is.

Interestingly, the academic sub-registers differ in their reliance on these two clause-connecting strategies, paralleling the differences shown for other grammatical complexity features. That is, as Figure 3.17 shows, overt linking adverbials are relatively common in humanities, social sciences, and popular science articles, but relatively rare in specialist science articles. In contrast, colons as clause connectors are much more common in social science and specialist science articles than in humanities or popular science. This pattern conforms to the overall trend in specialist science writing (and social science writing) to prefer maximally compressed structures, despite the associated loss in explicitness of meaning. These semantic relationships are discussed in detail in Chapter 6.

3.4 Chapter summary

The findings presented in the last section show that written academic discourse can be regarded as a microcosm of the ways in which grammatical complexity features vary in speech and writing generally. In [Section 3.3.1](#), we described how conversation differs from academic writing in the use of complexity features: conversation makes frequent use of verbs and finite dependent clauses and grammatical features that function syntactically as modifiers or complements of a verb. Academic writing makes frequent use of nouns, adjectives, and phrasal features that function as modifiers of a noun. In [Section 3.3.2](#), we discussed the characteristics of spoken classroom teaching, showing how that register is similar to conversation in its reliance on clausal structures. These findings indicate that production in the spoken mode is a more important determinant of grammatical structure than informational communicative purpose. [Section 3.3.3](#) then focused on written registers, showing how academic writing relies on phrasal noun modifying structures to a much greater extent than other ‘popular’ written registers (fiction and newspaper prose).

In the last section, we showed how similar patterns of variation exist among academic sub-registers. Thus, in comparison with other academic registers, humanities research writing patterns more like popular non-academic written registers in making frequent use of finite dependent clauses and clause types that function as modifiers of a verb. The largest difference among academic sub-registers is for the use of finite relative clauses, which are extremely common in humanities writing, but relatively rare in specialist science research articles. In contrast, specialist science and social science research writing are unlike virtually all other registers in their relative rarity of verbs and dependent clauses, and their extremely high frequencies of nouns and phrasal modifiers.

In sum, present-day academic research writing relies heavily on a wide array of phrasal grammatical devices. These features are structurally compressed rather than elaborated, with the result that they are not explicit in expressing the meaning relations among grammatical elements. These structures are generally common in academic writing: much more common than in any spoken register, and also more common than in most popular written registers. In addition, there are systematic patterns of variation among sub-registers within academic writing: humanities research writing is the most similar to popular written registers in its reliance on clausal modification; specialist science research writing is the

most distinctive in its extreme reliance on phrasal structures modifying nouns.

Thus, academic writing can be regarded as an outlier: a register unlike nearly all other spoken and written registers. Specialist science research writing is the most extreme representative of this style. We show in the following chapter that this style of discourse is a recent innovation, developing primarily in the twentieth century.

The historical evolution of phrasal discourse styles in academic writing

4.1 Grammatical complexity in eighteenth century written texts

In the [last chapter](#), we described the ways in which modern academic writing differs in its grammatical discourse style from most other registers of English, including all spoken registers as well as most other written registers. The findings in [Chapter 3](#) are surprising in two major respects: on the one hand, grammatical features associated with complexity in previous research turn out to be *not* frequent in academic writing, but at the same time, other features that have been overlooked in most previous studies turn out to be especially prevalent in academic writing.

For example, passive voice verbs have been strongly associated with academic writing in past research, often with a negative evaluation. Corpus research findings show that passive voice verbs are indeed more common in academic writing than in most other registers. However, it is not correct that passives are ubiquitous in academic writing. Rather, passives account for only c. 25% of all finite verbs in academic prose (see Biber et al. 1999: 476). They are reserved for specialized functions, often when it would be redundant to mention that the author is the agent of an action (e.g., in methodology sections of research articles). Similarly, although nominalizations are more common in academic prose than in other registers, they are not especially frequent in absolute terms: [Figures 3.8 and 3.11 in Chapter 3](#) show that nominalizations account for only 10%–15% of all nouns in written academic texts. And surprisingly, that proportion is actually lower in specialist science research articles than it is in the social sciences and humanities (see [Figure 3.11](#)).

The most surprising pattern documented in the [last chapter](#) is the relatively rare use of dependent clauses in written academic registers, contrasted with much more frequent use of finite clausal structures in spoken registers and popular written registers. Thus, finite adverbial clauses and complement clauses are considerably more common in speech than in academic writing. Finite relative clauses are commonly used in

humanities academic writing, but they are not common in science academic writing. Non-finite dependent clauses functioning as noun modifiers or noun complements have a stronger association with academic writing, yet even these structures are not especially frequent in absolute terms (see Figures 3.9 and 3.12).

So what makes academic writing complex, if most of those grammatical structures traditionally associated with complexity are not prevalent? We argued in the [last chapter](#) that the most important distinguishing characteristic of academic writing is its extremely dense use of phrasal structures (i.e., structures without verbs), especially phrases functioning as noun modifiers. These structures include attributive adjectives, nouns as nominal pre-modifiers, prepositional phrases as nominal post-modifiers, and appositive noun phrases. The dense use of these structures results in a discourse style that minimizes the use of verbs and clauses, relying instead on embedded phrases to convey information. These features have been mostly overlooked in previous research on academic writing, despite their quite frequent use in this register.

English academic writing has not always been this way. Thus, consider the following extended text sample taken from a scientific academic paper published in the *Philosophical Transactions of the Royal Society of London* in 1800. We have underlined all main verbs, and marked dependent clauses in [].

Text Sample 4.1

Eighteenth/nineteenth century academic science writing

It has long been my intention [to lay before the Royal Society a few observations on the subject of sound]; and I have endeavoured [to collect as much information], and [to make as many experiments, [connected with this inquiry], [as circumstances enabled me [to do]]; but, the further [I have proceeded], [the more widely the prospect of [what lay before me] has been extended]; and, [as I find [that the investigation, in all its magnitude, will occupy the leisure hours of some years, or perhaps of a life]], I am determined, in the mean time, [lest any unforeseen circumstances should prevent [my continuing the pursuit]], [to submit to the Society some conclusions [which I have already formed from the results of various experiments]].

Thomas Young, M. D. January 16, 1800.

‘Outlines of Experiments and Inquiries respecting Sound and Light’

F. R. S. In a Letter to Edward Whittaker Gray, M. D. Sec. R. S.

This text is typical of eighteenth and early nineteenth century science writing in its long, complex sentence structures, with multiple levels of clausal embedding. Both finite and non-finite dependent clauses are common in this discourse style, as shown by the bracketed clauses in [Text Sample 4.1](#). Embedded phrases are also relatively common, especially *of*-phrases (e.g., *the*

subject of sound; the prospect of . . . ; the leisure hours of some years, or perhaps of a life; the results of various experiments). However, many of the distinctive phrasal characteristics found in modern academic prose are relatively rare in early science texts, including prepositional noun modifiers other than *of*-phrases and nouns as nominal pre-modifiers. We would clearly characterize this style of writing as grammatically ‘complex’; but the complexity involves clausal embedding rather than the dense use of phrasal modifiers.

We show in the present chapter that the stereotypes about grammatical complexity in academic writing (with a dense use of dependent clauses) provide an accurate portrayal of science writing as it existed in earlier centuries, suggesting that this might be the source of those characterizations. However, as we have illustrated in detail in [Chapter 3](#), the eighteenth century discourse style illustrated in [Text Sample 4.1](#) is sharply distinguished from modern academic writing, and especially different from modern science research writing.

We further show in the present chapter that eighteenth century science writing was not especially distinctive in its use of complex clausal embedding. Rather, most other eighteenth century written registers were similar in employing a complex grammatical style with extensive use of dependent clauses. Even fiction tended to be structurally elaborated, employing many of the same kinds of structures as in [Text Sample 4.1](#). Thus consider the following passage from an eighteenth century novel written by Samuel Johnson:

Text Sample 4.2 Eighteenth century fiction

The valley, wide and fruitful, **supplied** its inhabitants with the necessities of life, and all delights and superfluities were **added** at the annual visit [which the Emperor **paid** his children, [when the iron gate was **opened** to the sound of music]], and, during eight days, everyone [that **resided** in the valley] was **required** [to **propose** [whatever might **contribute** [to **make** seclusion pleasant],] [to **fill up** the vacancies of attention, and **lessen** the tediousness of time]. Every desire was immediately **granted**. All the artificers of pleasure were **called** [to **gladden** the festivity]: the musicians **exerted** the power of harmony, and the dancers **shewed** their activity before the princes, in hope [that they should **pass** their lives in this blissful captivity, [to which those only were **admitted** [whose performance was **thought** able [to **add** novelty to luxury]]]].

Samuel Johnson. 1759.
‘Rasselas, Prince of Abyssinia’
Cassell and Company, Limited.

This relative lack of variation among written registers in earlier centuries has been documented in multi-dimensional (MD) studies of historical register

variation. For example, Biber and Finegan (1997/2001, see Figures 4 and 5) show that fiction and academic research writing were quite similar linguistically in the seventeenth and eighteenth centuries with respect to 'Dimension 1: Involved vs. Informational Production'. In contrast, these two registers are sharply distinguished from each other at present. Comparisons like this show that there has been extensive register diversification that has taken place over the last three centuries. The obvious question is how this change came about, and what societal forces may have influenced the course of change. We consider two of these influences in the following section: the increasing preference for colloquial forms in writing ('popularization'), and the need to efficiently convey a great deal of information ('economy').

4.1.1 *The influence of popularization versus economy in historical change*

One of most widely noted linguistic developments that has occurred in English over the past two centuries has been the increasing use of colloquial linguistic forms in written registers. As we discussed in [Chapter 1](#), researchers have given considerable attention to documenting the increased use of linguistic features associated with conversation in written registers. These features include first person pronouns, contractions, and semi-modals (e.g., *be going to*, *have to*). This trend, which has accelerated in the twentieth century, has been documented by discourse analysts and corpus linguists, and is referred to as the 'drift' of written registers towards more 'oral' styles (Biber and Finegan 1989a), 'informalization' (Fairclough 1992), and 'colloquialization' (Hundt and Mair 1999; Mair 2006; Leech et al. 2009).

Earlier investigations of these developments suggested that they represent a general historical trend affecting all written registers in English. However, subsequent investigations show that there are important differences among written registers, and as a result, some written registers have not fully participated in this historical shift. In particular, Biber and Finegan (1997/2001) show that written registers like fictional novels and personal letters have been strongly influenced by the shift to more colloquial linguistic styles, but written academic registers (science research articles and medical research articles) have not participated in these changes. Hundt and Mair (1999) also note this difference, distinguishing between 'agile' written registers (e.g., newspaper prose) that are receptive to these changes, versus 'up-tight' written registers (e.g., academic prose) that resist such changes.

However, it turns out that the actual patterns of historical change are more complicated than previously suspected. In particular, it is not accurate to describe academic prose as conservative and resistant to change. Rather, academic research writing has participated in alternative historical

linguistic changes, which are at least as important as the increased use of colloquial features documented for other written registers. These grammatical changes are completely different in kind. Thus, on the one hand it is accurate to portray academic writing as conservative because it has resisted the trend towards increased use of colloquial features. But at the same time, academic writing has been dynamic and innovative in its increased use of compressed structural devices, associated with ‘economy’ of expression (referred to as ‘densification’ in Leech et al. 2009).

The differences in the historical evolution of academic written registers versus other written registers can be interpreted relative to their underlying functional motivations, including social changes that have occurred over the past three centuries. On the one side are the influences of mass literacy and near-universal education, so that written texts are required for an increasingly wide readership. We refer to this influence as ‘popularization’ (see Biber 2003; Biber and Finegan 1989a). The general effect is to make writing accessible to the general public through adoption of a more informal style that employs features of conversation. At the same time, however, scientific research has been influenced by the opposite social force, so that academic sub-disciplines have proliferated and become increasingly specialized in topic, accessible only to restricted readerships. One underlying cause of this latter trend is the ‘information explosion’ and the need to present more information in an efficient and concise way; we refer to this factor as the need for greater ‘economy’ in written informational texts.

As we show in later sections, the linguistic innovations associated with economy in academic writing are at least as important as the shifts to colloquial styles witnessed in other written registers. The processes of popularization and colloquialization represent the spreading of existing linguistic patterns of use, from spoken discourse to writing. In contrast, we argue here that the historical developments towards increased economy in academic writing are more noteworthy, because they represent completely new linguistic styles of discourse not found in any register (spoken or written) in earlier historical periods.

The competing factors of popularization and economy have strongly influenced historical change in the grammatical characteristics of written texts over the past two centuries. As we show next, these factors correspond to different patterns of linguistic change: popularization is associated with the incorporation of colloquial features in written discourse, while economy is associated with the development of phrasal (rather than clausal) discourse styles. In particular, the historical processes associated with economy have affected the typical noun phrase structures found in

academic prose, and thus we turn to previous research on these structures in the following section.

4.1.2 *Previous historical research on the noun phrase in English*

One of the most salient linguistic characteristics of modern academic writing is its reliance on nominalizations and noun conversions. Nominalizations are nouns that have been morphologically derived from verbs (e.g., *development*, *progression*) or adjectives (e.g., *similarity*, *darkness*). ‘Conversions’ are verbs that come to be used as nouns, with no morphological additions (e.g., *increase*, *use*). Almost any sentence taken from a modern academic research article will illustrate the use of these features. For example, consider the opening sentences from three modern education research articles; derived nominalizations are **bold underlined** and converted nouns given in ***bold italics***:

Now that programed **instruction** has emerged from the laboratories of experimental psychologists and become a bona fide teaching ***aid***, **consideration** is being given to the **expansion** and **utilization** of the media by which programs are presented.

With the ***strain*** that increased **enrollments** are placing on the resources of many colleges and universities, these institutions are becoming more concerned with minimizing the **inefficiency** arising from student ***transfers***.

This paper reports an **analysis** of Tucker’s central **prediction** system model and an empirical **comparison** of it with two competing models.

In most cases, it is easy to restate this information using verbs and clauses. For example:

someone is considering the way in which the media has expanded and is being utilized

students are transferring from colleges and enrolling in other colleges

we analyzed a model that predicts systems, and compared it to other models

However, in modern academic writing, it is common to use nominalized rather than verbal structures to express information of this type.

Halliday (2004: xvi–xvii) refers to this characteristic of academic writing as a heavy reliance on ‘grammatical metaphor’, where nominalizations are used to refer to the normal meanings associated with verbs (‘processes’) rather than the normal meanings associated with nouns (‘entities’). Halliday (2004) provides examples to illustrate how discourse in written scientific prose exploits grammatical metaphor. For example, he contrasts alternative styles of expression, as in the following examples (Halliday 2004: 34):

1. *Glass cracks more quickly the harder you press on it.*
2. *Cracks in glass grow faster the more pressure is put on.*
3. *Glass crack growth is faster if greater stress is applied.*
4. *The rate of glass crack growth depends on the magnitude of the applied stress.*
5. *Glass crack growth rate is associated with applied stress magnitude.*

Halliday describes the first of these examples as the most ‘congruent’, where the meanings of words correspond to the expected meanings of the grammatical categories used. For example, the verbs *cracks* and *press* are used to refer to processes. In contrast, examples 4 and 5 illustrate a dense use of grammatical metaphor, with qualities and processes being expressed by nouns rather than adjectives and verbs. Halliday (1988; reprinted in Halliday 2004: chapter 5) has further claimed that science prose followed a historical progression similar to that illustrated in sentences 1–5 above. Illustrated with historical scientific texts written by Chaucer (1390), Newton (1704), and Priestley (1760s) – in contrast to modern science texts – Halliday argues that science prose has shifted historically from ‘congruent’ styles of expression to a dense reliance on ‘grammatical metaphor’.

Although they have been influential, Halliday’s contributions in this area have been primarily theoretical rather than empirical. That is, they have been fairly limited in the scope of the grammatical features considered, and they are not supported by a large-scale empirical analysis of historical texts. To address some of these concerns, Banks (2008) applies this theoretical framework to an empirical analysis of thirty science articles published in the *Philosophical Transactions of the Royal Society of London*. Banks’s study shows that nominalizations have increased in use historically, in both the physical and the biological sciences (see Banks 2008: 124). Most of these nominalizations are used to refer to ‘material processes’ (e.g., *separation*, *emergence*), but ‘mental processes’ (e.g., *calculation*) and ‘verbal [i.e., communicative] processes’ (e.g., *description*) are also important (see Banks 2008: 125–130).

Halliday and Banks focus almost entirely on a single grammatical feature in academic writing: nominalizations. However, there is actually a suite of related grammatical features that are similar in that they present information through ‘compressed’ phrasal structures rather than through the use of verbs and clauses. For the most part, these features are noun phrase modifiers, and several recent studies have documented strong historical increases in the use of this wider array of grammatical features in academic writing (see, e.g., Biber and Clark 2002; Biber 2009; Biber and Conrad 2009; Biber and Gray 2010, 2013b; Leech et al. 2009). Five grammatical devices have been especially important in these discussions:

- nominalizations (e.g., *consumption*, *comparison*, *sustenance*),
- attributive adjectives (*gradually expanding cumulative effect*),
- nouns as nominal pre-modifiers (*baggage inspection procedures*),
- prepositional phrases as nominal post-modifiers (*a high incidence of heavy alcohol consumption amongst patients*),
- appositive noun phrases (*Dallas Salisbury, CEO of the Employee Benefit Research Institute*)

As noted earlier, nominalizations have been the most studied of these five devices. However, historical change in the use of the other grammatical devices is actually more noteworthy, because they represent features which are both more common overall, and which have witnessed much more dramatic increases in frequency over time. For example, we show that the historical increase for nouns as nominal pre-modifiers over the course of the twentieth century is much more important in these respects than the increased use of nominalizations.

These historical developments represent a fundamental shift in the discourse style of academic writing. These changes have resulted in grammatical patterns of use that are not attested in any register in earlier historical periods. However, they entail considerably more than just stylistic shifts in the density of nominal/phrasal features. Rather, the individual grammatical features have also undergone important extensions in their uses and functions. In the following sections and chapters, we explore these changes in detail.

4.2 General patterns of change in written English registers

Similar to the synchronic descriptions in [Chapter 3](#), it is useful to employ a comparative register approach to highlight the distinctive historical patterns of change for academic research writing. That is, by comparing the change of grammatical characteristics in academic writing to the patterns of change in other general registers, it is possible to highlight the distinctive historical trajectory of that register. The present section thus compares the historical development of three general registers: fictional novels, news reportage (newspapers and magazines), and science research writing.

Historical change in the linguistic characteristics of written registers can be interpreted relative to the competing functional forces introduced in [Section 4.1.1](#). On the one side are the influences of mass literacy and near-universal education. Because general literacy rates were low in the seveneenth and eighteenth centuries, written texts were addressed to a small readership. By the twentieth century, though, most adults in western

countries came to participate in formal education and thus developed relatively high standards of literacy. As a result, written registers like novels, newspapers, and magazines evolved to adopt linguistic styles that are accessible to this increasingly wide readership. In contrast, academic research writing has been influenced by the opposite forces over the last two centuries: the domain of academic research increased dramatically, accompanied by a proliferation of sub-disciplines that became increasingly specialized in both topic and readership. This ‘information explosion’ resulted in the need to present information in an efficient and concise way. As the present section shows, both of these major factors have influenced the linguistic development of English written registers over the last three centuries.

To track these patterns of change, we grouped texts into six general time periods: 1700–1800 (labeled ‘1750’), 1801–1870 (labeled ‘1850’), 1871–1919 (labeled ‘1900’), 1920–1949 (labeled ‘1930’), 1950–1979 (labeled ‘1960’), and 1980–2005 (labeled ‘1990’). We are interested primarily in change over the last century, and so we undertook more detailed sampling for those periods (see [Chapter 2](#) for additional details on corpus design and composition).

[Appendix 1](#) provides detailed descriptive statistics for several major grammatical characteristics, giving means and standard deviations for each linguistic feature in each register, across these historical periods. However, it can be difficult to determine the overall patterns of change from these detailed statistics.

Pearson correlations provide a more useful statistic for determining general historical trends. In the present case, we correlated the rate of occurrence for a linguistic feature in each text with the year of publication for those texts. [Figure 4.1](#) illustrates the data used to compute these correlations, plotting the rate of occurrence for nouns in science research articles versus the year of publication (grouped into general historical periods) for each text.

Although there is variation among the texts within each historical period, [Figure 4.1](#) shows a strong overall trend for an increased use of nouns in science writing across these historical periods. That trend corresponds to a strong correlation: $r = .77$. That is, correlations determine whether there is an association between the rate of occurrence of a linguistic feature and historical period, measuring both the extent and the direction of change for a linguistic feature in a register. Correlation coefficients have a scale of -1 to $+1$: a value near -1 represents a strong decrease in use over time; a value near $+1$ represents a strong increase in use over time; and a value near 0.0 indicates that there has not been any consistent pattern of change over time.

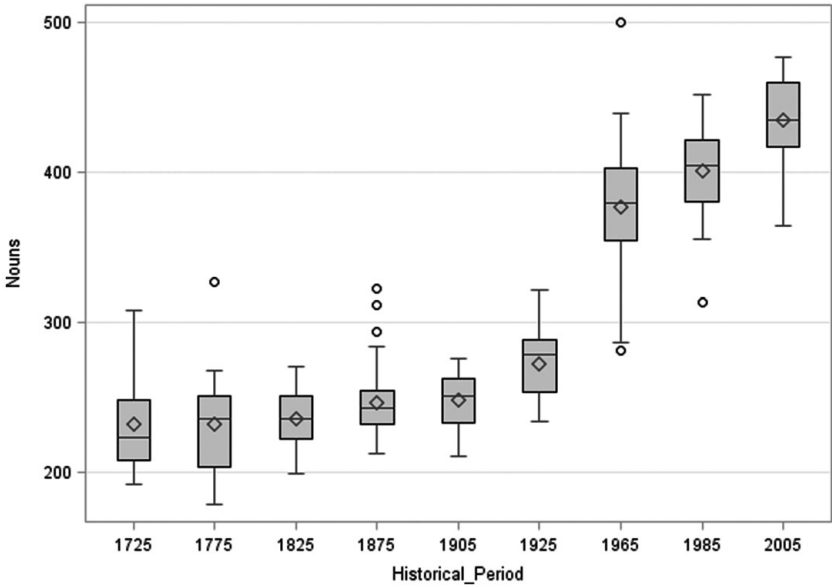


Figure 4.1 Distribution of nouns (per 1,000 words) across historical periods

Table 4.1 presents Pearson correlation coefficients for eight general linguistic characteristics, correlating the rate of occurrence for those features with historical period. The general historical trend shown by the correlation coefficient is also indicated: moderate increases in use are shown by + and stronger increases are shown by ++ or +++; decreases in use are shown by - and --.

Table 4.1 shows that each register has undergone distinct patterns of historical development over the past three centuries. In fiction, lexical verbs and adverbs have increased in use, while passive voice verbs, nominalizations, prepositional phrases, and relative clauses have all decreased in use. Overall, words have become shorter (corresponding to the large decrease in the use of nominalizations).

Science research writing has developed in exactly the opposite ways: lexical verbs and adverbs have decreased strongly over time, while nouns, adjectives, and nominalizations have increased in use. Words, on average, have become much longer over time in this register.

News reportage is intermediate between these two opposing trends, incorporating elements of both. Thus, lexical verbs and adverbs have increased in use (similar to fiction), but nouns and adjectives have also increased in use, and words have become longer overall (similar to academic writing).

Table 4.1 *Historical change in the use of general linguistic features, shown by Pearson correlation coefficients (r) for the rate of occurrence correlated with date (1750–1990)*

	Fiction N = 215 texts		News reportage N = 1140 texts		Science prose N = 524 texts	
	trend	r	trend	r	trend	r
Nouns		.17	++	.34	+++	.77
Lexical verbs	++	.33	+	.28	--	-.55
Adjectives		.01	++	.31	++	.37
Adverbs	++	.47	+	.25	--	-.58
Passive verbs	--	-.38	--	-.44		-.07
Nominalizations	--	-.55		.04	++	.48
Prepositional phrases	--	-.40	--	-.55	--	-.42
Average word length	--	-.32	+	.29	+++	.60

Key:

.60 to .99 = +++

.30 to .59 = ++

.20 to .29 = +

-.20 to -.29 = -

-.30 to -.59 = --

The increase in use for nouns in science prose is the strongest change shown in Table 4.1 (see also Figure 4.1). In contrast, fiction shows little change over time in the use of nouns (reflected by the correlation of only .17), while news reportage shows a moderate increase ($r = .34$). Nominalizations have also increased in science prose, but not nearly to the same extent as simple nouns (see Section 4.2.3). The frequency of nominalizations has remained essentially unchanged in news reportage, while there has been a strong decrease in use in fiction.

The historical trend for prepositional phrases – decreasing in all registers – requires further explanation. This trend appears to contradict our findings in Chapter 3, which singled out prepositional phrases as an important distinctive characteristic of modern academic research writing. The correlations reported in Table 4.1 are based on analysis of all prepositional phrases, including phrases functioning as adverbials and phrases functioning as noun modifiers. These phrases include both *of*-genitives and phrases with other prepositions. The findings in Chapter 3, however, show that the different types and functions of prepositional phrases have different register distributions. More specifically, the analysis in Chapter 3 showed that:

- Prepositional phrases functioning as adverbials are actually more common in modern fiction and news reportage than they are in modern academic prose (see [Table 3.8](#)).
- In contrast, prepositional phrases functioning as noun modifiers are considerably more common in modern academic prose than in other present-day written registers (see [Table 3.8](#)).
- There are complementary distributions for *of*-phrases versus other prepositions as noun modifiers across disciplines: *of*-phrases are more common in humanities academic prose than in the sciences, while other prepositional phrases are considerably more common as noun modifiers in the sciences (and social sciences) than in other disciplines (see [Table 3.11](#)).

It turns out that the overall decrease in use for prepositional phrases (shown by the negative Pearson correlations in [Table 4.1](#)) hides more interpretable patterns of increase and decrease for the specific types and syntactic functions. We thus return to a more detailed consideration of prepositional phrases in [Section 4.2.3](#).

Some of the historical trends in [Table 4.1](#) are shared across registers. For example, words have become longer on average, especially in academic science prose, but also in news reportage. Passive voice verbs have decreased in use in all registers. This trend is strongest in fiction and news reportage. Surprisingly, the frequency of passive voice verbs has also decreased in academic science writing, although there has been only a slight change overall. This trend might be related to the more general drift towards colloquial styles, especially in fiction and news reportage. We thus turn to more detailed consideration of colloquial features in the following section.

4.2.1 *Increasing use of colloquial features*

As noted earlier, one of the most noticeable linguistic developments that has occurred in English over the past two centuries has been the increasing use of colloquial linguistic forms in written registers, such as contractions, semi-modals (e.g., *be going to*, *have to*), and even first person pronouns. In the last section, we documented historical trends for general grammatical characteristics associated with written discourse. In the present section, we shift our attention to the use of these colloquial features, demonstrating that the popularization of writing has had a major influence on general registers like fiction and news reportage. However, the corpus findings presented next show that academic prose has been little influenced by this trend.

[Table 4.2](#) presents Pearson correlation coefficients for many of the most important colloquial features discussed in previous research, such as

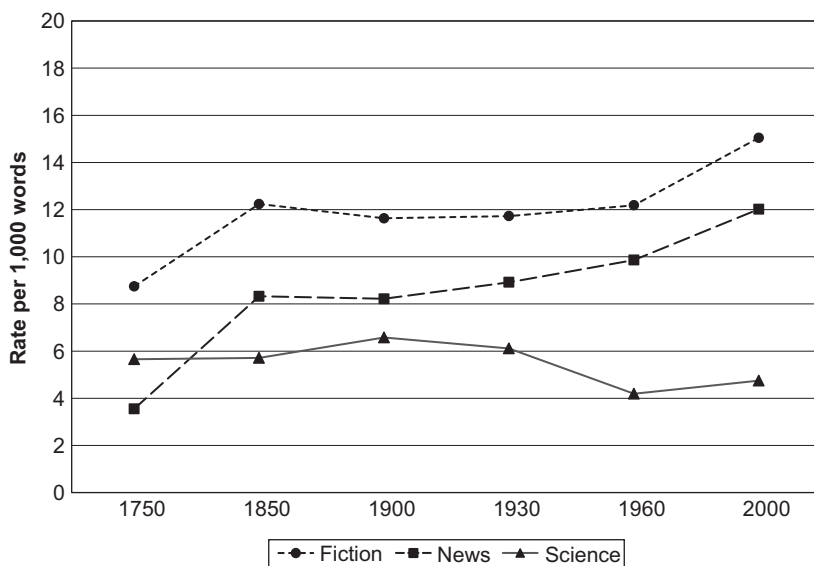


Figure 4.2 Historical change in the use of progressive aspect verbs

decreased in use ($r = -.44$). Similarly, bare infinitive constructions increased in use, while *to*-infinitives remained relatively unchanged in use over time.

As noted earlier, many of these features have also increased in use in news reportage, but the changes are much less pronounced than in fiction. Progressive aspect verbs provide a clear example of colloquialization affecting both fiction and news reportage (see also Leech et al. 2009: chapter 6). Figure 4.2 shows that news reportage lagged behind fiction in the use of progressives across centuries, but both registers follow the same general course, with an especially notable increase in use during the second half of the twentieth century. In contrast, science prose has changed little over the centuries in the use of progressive verbs, with a slight decrease in the twentieth century.

Table 4.2 further shows that complex progressives (i.e., perfect-progressives and progressive-passives) have generally increased in use. The overall frequencies for perfect-progressives are much lower than for simple progressives, but the historical development is quite similar: a gradual increase across the centuries, and a stronger increase during the twentieth century, with news reportage lagging behind fiction. Thus, perfect-progressives are moderately common in modern-day fiction and news reportage; for example:

Fiction:

People had been going about their business past those walls year after year.

News:

Richard Ramirez is a native of El Paso, Texas, who has been drifting around Los Angeles and San Francisco for several years.

Progressive-passives are a more interesting case, because their historical trend is the opposite of that for passive voice verbs generally. That is, in [Section 4.2.1](#), we described how passive voice verbs have decreased in use in the twentieth century, especially in news reportage, but also in fiction and even in academic science prose. This decrease might be interpreted as a by-product of the overall trend towards increased colloquialization, by arguing that it reflects a dispreference for this stereotypically 'literate' feature.

However, in contrast to that general trend, progressive-passive verb phrases have increased in use in all three registers (although the actual frequencies are extremely low: less than 0.2 per 1,000 words). This trend can be interpreted as grammatical change in the use of progressive aspect, because it is recently coming to be used with passive voice. These verb phrases can occur in both finite and non-finite clauses; for example:

Finite clauses:

Fiction:

As a matter of fact, he was being punished.

News:

Policy papers are being prepared for the President.

Academic:

Three analytical approaches are being compared with the simulated data sets.

Non-finite clauses following a preposition:

Fiction:

They were instinctively aware of being watched.

News:

The looming threat of being kidnapped and murdered is crippling his ability to launch an adequate defense of the fallen dictator.

Academic:

Local assemblages within a region can show significant levels of nestedness despite being dominated by taxa with good powers of inter-locality dispersal.

Finally, first and second person pronouns have remained essentially unchanged over time in fiction and news reportage. Both types of pronouns were already common in fiction in the eighteenth century, and they have remained common up to the present time. In eighteenth century news reportage, first person pronouns were moderately common, while second person pronouns were rarely used; that pattern of use also persists up to the present day with little change. In contrast, academic prose has shown a marked decrease in the use of these pronouns. In the eighteenth century, first person pronouns were moderately common in academic writing, but they have become considerably less common in present-day academic texts. Second person pronouns were already rare in eighteenth century academic texts, and they have become even less frequently used over time.

In summary, the influence of colloquialization is strongest in popular written registers: written fiction has changed over the centuries to adopt an increased use of numerous conversational features. News reportage – a popular register with an informational focus – has undergone smaller changes towards increased use of these colloquial features. The strongest patterns for news reportage are for progressive verbs and passive voice verbs: progressives have increased notably in news reportage over the last century, while passive voice verbs have shown a strong decrease in use (compare [Tables 4.1](#) and [4.2](#)).

In contrast to the historical development of popular written registers, colloquialization has had little influence on the discourse style of academic prose, and thus these colloquial features are rarely used at present in that register. Overall, it is clear that colloquialization has been a major trend influencing historical change in written discourse over the past three centuries, but it has not applied uniformly to all written registers. As the following sections show, the opposing trend towards increased structural compression has been at least as important as a functional force influencing historical change in written discourse. However, that influence has operated in a complementary manner, affecting written academic registers to a much greater extent than popular written registers.

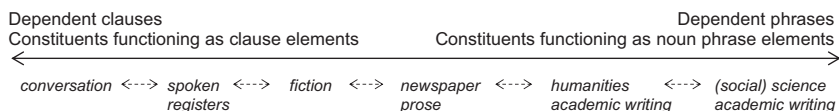
4.2.2 Historical change in the use of complexity features

In [Chapter 3](#), we discussed the use of grammatical features associated with increased ‘complexity’ in various spoken and written registers. Two major grammatical oppositions were explored in that chapter:

- dependent clauses versus dependent phrases
- syntactic role as clause constituent/modifier versus noun modifier

The two parameters tend to go together when one is considering the grammatical structures that are most frequent in conversation versus academic writing. That is, conversation employs extensive clausal complexity, realized as a dense use of dependent clauses functioning as clause constituents (complement clauses) or clausal modifiers (adverbial clauses). In contrast, modern academic writing employs extensive phrasal complexity, realized as a dense use of embedded phrases functioning mostly as modifiers of a head noun.

In [Chapter 3](#), we compared the use of these complexity devices across a range of present-day spoken and written registers. Those analyses show that it is not accurate to simply characterize a register as more or less complex. Rather, different registers employ different kinds of grammatical complexity. The most surprising finding from [Chapter 3](#) is that clausal complexity – the focus of most previous research – is actually more characteristic of conversation and spoken registers than written registers. In contrast, the complexity of informational writing resides primarily in a dense use of embedded phrasal structures. Thus, the synchronic register comparisons from [Chapter 3](#) can be summarized as a cline ranging from clausal to phrasal discourse styles:



The research question to be explored in the present section is how these different discourse styles developed historically. [Table 4.3](#) presents Pearson correlations that measure historical change in the use of several of these complexity features in three general registers (fiction, news reportage, and academic prose).

The density of finite dependent clauses functioning as clause constituents has changed little in these registers over the past three centuries. Adverbial clauses and complement clauses have increased slightly in fiction and news reportage, and decreased slightly in academic prose (see the descriptive statistics in [Appendix 1](#)). However, the correlations reported

Table 4.3 *Historical change in the use of complexity features, shown by Pearson correlation coefficients (r) for the rate of occurrence correlated with date (1750–1990)*

	Fiction N = 215 texts		News reportage N = 1140 texts		Science prose N = 524 texts	
	trend	r	trend	r	trend	r
Dependent clauses						
Finite adverbial clauses		.13	+	.24	-	-.22
Finite complement clauses		.16		.12		-.15
Relative clauses (total)	--	-.47		-.05	--	-.61
NP clausal post-modifiers						
WH-relative clauses	--	-.50	--	-.35	---	-.62
That relative clauses		.02	++	.42		-.06
Non-finite -ing relatives	--	-.36	--	-.38	--	-.36
Non-finite -ed relatives		.06		-.15		.09
That noun complement clauses	+	.25		.04	+	.24
To noun complement clauses		-.13		.14		-.11
Of + -ing clause		-.12		-.10		-.12
Other preposition + -ing clause		-.09		-.01		.16
NP phrasal pre-modifiers						
attributive adjectives	-	-.25		.09	+	.29
Nouns as nominal pre-modifiers	++	.51	++	.53	+++	.75
N-xxing as pre-modifiers	+	.22		.12	+	.28
N-xxed as pre-modifiers	+	.26	+	.21	++	.33
NP phrasal post-modifiers*						
of-phrases as post-modifiers	--	-.47	--	-.57	-	-.20
in-phrases as post-modifiers	+	.26		.17	+++	.60
on-phrases as post-modifiers	++	.43	+	.20	++	.41
for-phrases as post-modifiers		-.10		.11	+++	.60
with-phrases as post-modifiers		-.02	+	.29		.13

Key:

.60 to .99 = +++

.30 to .59 = ++

.20 to .29 = +

-.20 to -.29 = -

-.30 to -.59 = --

-.60 to -.99 = ---

* These findings regarding the distribution of prepositional phrases are based on hand-coding of twenty texts from each sub-register for 1850–1985. We identified only prepositions that immediately followed a noun, considering all occurrences for less common prepositions (*for* and *with*) and sampling from across the text for more common prepositions (*in* and *on*). Then, each prepositional phrase was coded by hand to determine whether it was functioning as an adverbial or noun modifier. The rates of occurrence for each text were then extrapolated, based on the overall counts for those prepositions in the texts.

in Table 4.3 are small, showing that there has been little systematic change in the use of those clause types.¹

In contrast, our results show a relatively strong decrease in the use of relative clauses, especially in fiction and academic research writing. (The overall frequency of relative clauses in news reportage has remained essentially unchanged.) This finding is surprising, because relative clauses have been regarded as one of the major indicators of complex academic writing. Upon closer inspection, it becomes clear that the overall decrease in use is due mostly to a large decline in the frequency of *WH*-relative clauses (see also Hinrichs et al. to appear). *That* relative clauses have increased moderately in news reportage (especially in American English), but they are less frequent than *WH*-relative clauses in absolute terms, and their density of use has remained unchanged in fiction and academic prose. Non-finite relative clauses show a similar pattern: *-ing* clauses have decreased in use in all three registers, while *-ed* clauses have remained essentially unchanged in their density of use. Noun complement clauses have also changed little in their overall density of use: *that* noun complement clauses have increased slightly in fiction and academic prose, but *to*-clauses and *-ing* clauses have remained essentially unchanged in their density of use. Taken together, these trends reflect an overall decline in the reliance on clausal subordination; this decline has been by far most pronounced in academic prose, but it has to some extent affected all written registers.

Thus, written discourse has certainly not become more ‘complex’ over the last three centuries in terms of its reliance on dependent clause structures. In fact, to the extent that there has been change in the use of these features, the trend has been towards decreased use. To some extent, that trend has influenced all written registers, but it has been strongest in academic writing.

In contrast, phrasal modifiers of noun phrases have generally increased in use, with very strong increases found for some features; those increases have been most pronounced in academic research writing. This pattern

¹ There have, however, been important changes in sentence style over the past three centuries, with sentences in eighteenth century prose often being very long, with multiple dependent clauses. Further, eighteenth century written prose differed from twentieth century prose by incorporating the full inventory of dependent clause types, including finite and non-finite clauses in all syntactic functions. Thus, while the overall density of dependent clauses in texts has changed little over the past three centuries, there has been a striking evolution in sentence style (contrast text Samples 4.1 and 4.2 with the text samples found in Chapter 3).

exists for both nominal pre-modifiers and nominal post-modifiers. Among the pre-modifiers, nouns as nominal pre-modifiers were relatively rare in the nineteenth century, but they have steadily increased in use over the subsequent decades (see [Section 4.2.3](#)). This change occurred in all three written registers, but there has been an especially large increase in academic prose (shown by the large positive correlation of .75). Pre-modifying nouns that are further post-modified by a participle (e.g., *law-making powers*, *age-related change*) are not frequent in absolute terms, but they have increased in use in all registers. Attributive adjectives are the only exception to the increased use of phrasal noun pre-modifiers: this feature was already common in all three written registers in the nineteenth century, and it has continued to be common (i.e., with little change in frequency) up to the present day.

In contrast to the decline seen for relative clause constructions, phrasal post-modifiers of a head noun have generally increased in use over the last three centuries. That increase has been especially strong in academic prose, but [Table 4.3](#) shows that it has also occurred to a lesser extent in fiction and news reportage. The affected features here include prepositional phrases as noun modifiers as well as appositive noun phrases (see [Figure 4.10](#)).

The one exception to this general trend is post-nominal *of*-phrases. Similar to attributive adjectives, *of*-genitive phrases were already well established and frequent by the eighteenth century, and they have not increased in use since then. In fact, this structure decreased in use in all three written registers over the last century (although *of*-phrases are still common in academic writing – see [4.2.3](#)).

In contrast, other prepositions as post-nominal modifiers were rare in the nineteenth century, but have shown strong increases in use, especially during the twentieth century. [Table 4.3](#) shows the correlations for *in*, *on*, *for*, and *with* phrases. The increases in use are strong in academic writing (see especially the large positive correlations for *in*, *on*, and *for* phrases), and moderately strong in news reportage. However, even fiction shows relatively large increases in use for *in* and *on* as post-nominal modifiers.

In summary, two major patterns are evident from [Table 4.3](#). First, the most important historical changes for complexity features have occurred with phrasal rather than clausal grammatical features. Phrasal noun modifiers that were already well-established by 1800 – attributive adjectives and post-nominal *of*-phrases – did not participate in this historical trend

(and *of*-phrases have actually decreased in use.) But many other types of phrasal noun modifiers were just becoming established in the eighteenth century, and these features have all undergone large increases in use over the subsequent decades.

Second, although these historical changes have affected all written registers, they have been most important in the informational registers; the changes have been most pronounced in academic writing, with news reportage also undergoing large increases in the use of these features. The most important functional motivation for these changes is apparently an informational purpose (characteristic to some extent of both news reportage and academic writing), with an expert/specialized readership being a secondary factor (in academic writing). In addition, some of these changes that were initiated in informational writing have subsequently been adopted in other written registers; thus, popular written registers like fiction have undergone many of these changes to a lesser extent. However, as shown in [Chapter 3](#), these features remain especially characteristic of informational written discourse, while all of these features are extremely rare in present-day conversation and other spoken registers.

4.2.3 *Exploring the magnitude of change: The evolution of a phrasal discourse style*

The preceding sections use correlation coefficients to track systematic linguistic changes across historical periods. Specifically, we computed rates of occurrence for each linguistic feature in each text, and then used Pearson correlations to measure the extent to which there have been systematic increases/decreases in those rates across historical periods. Those analyses show that phrasal structural devices functioning as noun phrase modifiers have increased in use over time, while dependent clauses have declined in use, even when they function as noun modifiers (i.e., relative clauses).

One limitation of correlational statistics is that they do not indicate the overall frequency of a feature or the magnitude of change. So, for example, an increase from 2 to 4 to 6 occurrences (per 1,000 words) across historical periods would result in a similar positive correlation as an increase from 100 to 200 to 300 occurrences (per 1,000 words) across periods. Similarly, a larger correlation for a feature in a register does not necessarily mean that the feature is more frequent in that register. So, for example, an increase

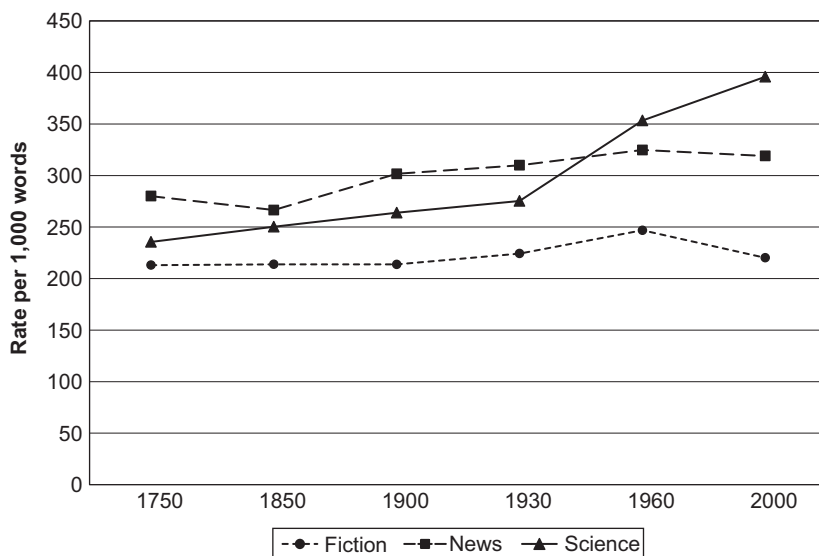


Figure 4.3 Historical change in the use of nouns

from 2 to 4 to 6 occurrences (per 1,000 words) across historical periods would result in a positive correlation in Register A, while a decrease from 300 to 200 to 100 occurrences (per 1,000 words) would result in a negative correlation in Register B – even though the actual rate of occurrence for the feature is much higher in Register B. Thus, from a textual perspective, the overall frequency and the magnitude of change must also be considered to fully document historical developments.

A second limitation of Pearson correlations for the analysis of historical change is that they capture linear developments. However, the actual pattern of change for many of these features has not been linear. Rather, there have been gradual patterns of change across the eighteenth and nineteenth centuries, followed by much more dramatic shifts in the patterns of use in the twentieth century.

For example, Figure 4.3 plots historical change in the density of common nouns in three general registers, corresponding to the correlation coefficients given in Table 4.1. Figure 4.3 shows that nouns are relatively common in fiction, but their rate of occurrence has changed little over these historical periods ($r = .17$ in Table 4.1). Nouns are more frequent overall in news reportage, and there has been a more consistent increase in use over time, although that change has not been especially strong ($r = .34$ in Table 4.1).

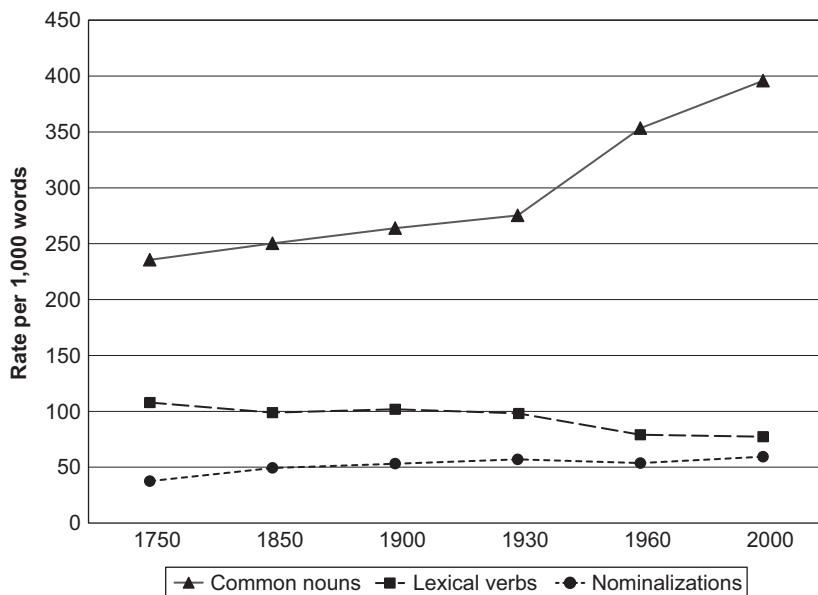


Figure 4.4 Historical change in science research writing: common nouns, nominalizations, and verbs

In contrast, we can see from [Figure 4.3](#) that nouns were actually less frequent in science writing than in news reportage during the eighteenth and nineteenth centuries. However, that relationship changed dramatically over the course of the twentieth century. The historical change in the use of nouns in academic science writing is non-linear: a gradual increase in use during the eighteenth and nineteenth centuries, followed by a rapid increase in use during the twentieth century. As a result, nouns are currently much more frequent in science writing than in either of these other two registers.

We noted in [Chapter 3](#) that nominalizations have been singled out as one of the most distinctive linguistic characteristics of scientific writing (see, e.g., [Halliday 2004](#); [Banks 2008](#)). However, [Figure 4.4](#) shows that common nouns are much more frequent than nominalizations in science prose. In addition, [Figure 4.4](#) shows that nominalizations have undergone only a gradual linear increase in use over time, while common nouns have rapidly accelerated in use during the twentieth century. As a result, common nouns are eight times more frequent than nominalizations in present-day science research writing. ([Figure 4.3](#) also shows that lexical

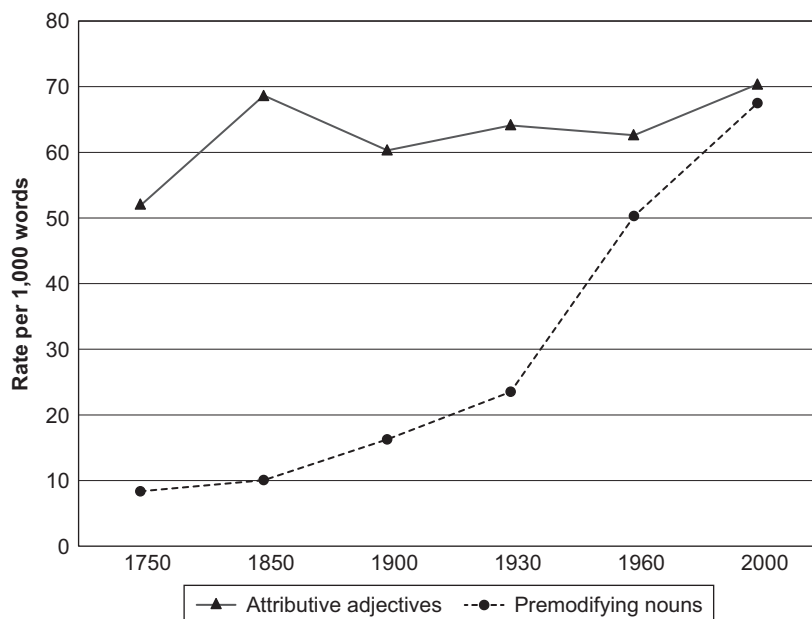


Figure 4.5 Historical change in the use of noun phrase pre-modifier types in academic prose: attributive adjectives versus nouns as pre-modifiers

verbs have been gradually decreasing in use in academic research writing over the past three centuries; we further discuss this trend next.)

In modern science writing, many of these nouns function as nominal pre-modifiers (e.g., *sea level change*). However, Figure 4.5 shows that was not always the case. In fact, pre-modifying nouns were rare up until the twentieth century. However, at that point there was a large increase in use in academic prose. The historical shift begins at the turn of the twentieth century, and then rapidly takes off in the mid-twentieth century, continuing to increase in use right up to the present time. In contrast, attributive adjectives were already common as noun modifiers in the eighteenth century, and they have continued to be common up to the present day, with little change.

In Chapter 3, we described the use of a highly specialized structure found in late-twentieth-century academic research writing: noun-participle compounds functioning as nominal pre-modifiers (e.g., *accountability-based reform efforts*, *decision-making practices*). As Figure 4.6 shows, this structural device was rarely used throughout the eighteenth and nineteenth centuries, with little change in use even through the first half of the

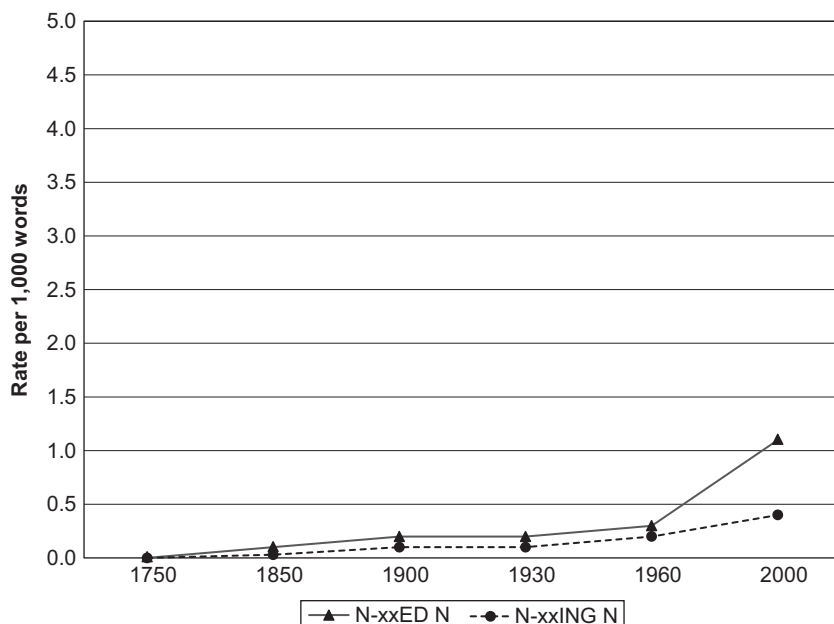


Figure 4.6 Historical change in the use of emerging NP pre-modifier types in academic prose: N-XXing + N and N-XXed + N

twentieth century. However, in recent decades, these devices have come to be used with increasing frequency. This is especially true for N-*ed* + N constructions, which have increased in use by a factor of 4 over the last 50 years. However, N-*ing* + N constructions have also more than doubled in that period. Neither of these features is especially frequent in absolute terms, but they are both showing rapid recent increase in use, with no signs of a leveling off in those trends.

Two other structural devices that function as post-nominal modifiers have actually decreased in use over the past three centuries: relative clauses and *of*-phrases. Figure 4.7 shows that relative clauses were common in all written registers in the eighteenth century, reflecting the generally elaborated style of written discourse typical of that period (see Section 4.1). The frequency of use declined during the nineteenth century, and then remained relatively stable up through the early twentieth century. Over those historical periods, relative clauses were moderately frequent in news reportage and academic prose, while fiction evolved to employ a less elaborated style, with fewer relative clause constructions. In the second half of the twentieth century,

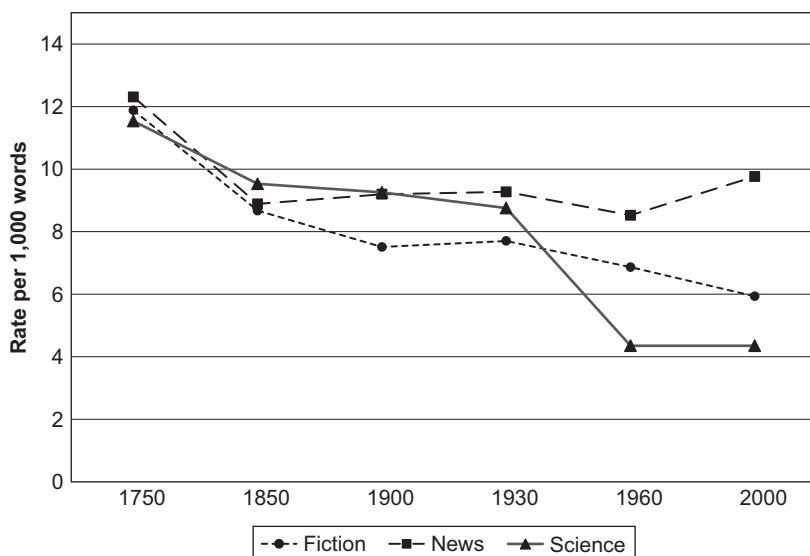


Figure 4.7 Historical change in the use of finite relative clauses

relative clauses continued to decline gradually in fiction (reflecting the general shift towards more colloquial styles), while they have remained essentially constant in use in news reportage. However, the most notable change here occurred in academic research writing, which has undergone a dramatic reduction in the use of relative clauses over the last fifty years.

We noted in the discussion of Table 4.2 the surprising historical decline for prepositional phrases in academic prose. However, there are different structural types of prepositional phrases (especially the *of*-genitive versus other prepositional phrases), and those structures can serve adverbial versus noun-modifying syntactic functions. Figure 4.8 shows that these grammatical differences are important for understanding the actual patterns of historical change: in science writing, there has been a sharp decrease in use for prepositional phrases functioning as adverbials, and a similarly strong decline for *of*-phrases functioning as noun modifiers; however, there has been an equally strong *increase* in use for other prepositional phrases functioning as noun modifiers. As a result, sentences like the following are common in present-day academic prose:

Specifically, we were interested in the qualitative ecological difference in emphasis between changes in composition vs. changes in relative abundance.

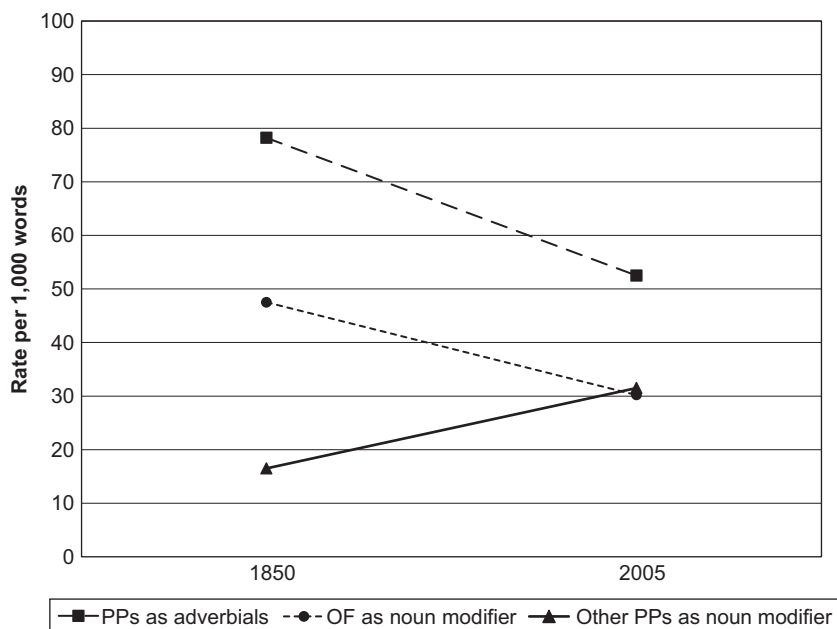


Figure 4.8 Historical change in the distribution and functions of prepositional phrases in science academic prose

Figure 4.9 traces the development of four individual prepositions functioning as noun modifiers: *in*, *on*, *for*, *with*. Although none of these prepositions is especially frequent by itself, they have all been increasing steadily over the last 100 years. Similar increases can be observed for many other prepositions that head post-nominal phrases. Taken together, these ‘other’ prepositional phrases functioning as noun modifiers have become as common as *of*-phrases in modern science prose. However, there are signs that this historical trend is leveling off, especially when considered relative to the full suite of other structural devices used for nominal modification. We return to that comparative perspective in Chapter 5.

Appositive noun phrases are another phrasal device used for noun modification. In Chapter 3, we discussed the frequent use of this structural device in twentieth century news reportage and academic research writing. Historically, these structures date back to Middle English (see Nevalinna and Pahta 1997; Pahta and Nevalinna 1997), but they were relatively rare in the eighteenth and nineteenth centuries. However, as Figure 4.10 shows, appositive noun phrases have become considerably more common in the twentieth century.

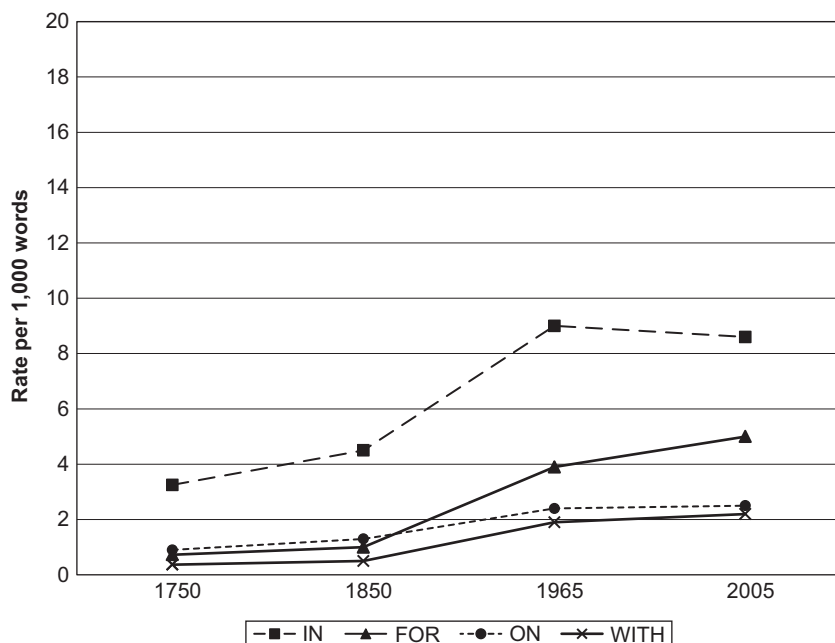


Figure 4.9 Historical change in the distribution of specific prepositional phrases functioning as noun modifiers in science academic prose

The typical function of appositive noun phrases in earlier centuries was to add descriptive information about a human head noun. This function was common in both news reportage and academic prose, and it continues to be one of the primary uses in twentieth century news reportage:

Eighteenth century newspaper:

About ten days ago Sir George Benham, **the Governor of Hong Kong**, received a message...

Eighteenth century science prose:

A post-mortem examination was made by Dr. C.W. Swan, **pathologist to the city hospital**...

In the nineteenth century, a structural/textual innovation was introduced in science writing: the use of parentheses to mark the appositive noun phrase, rather than separating the two noun phrases by commas. For example:

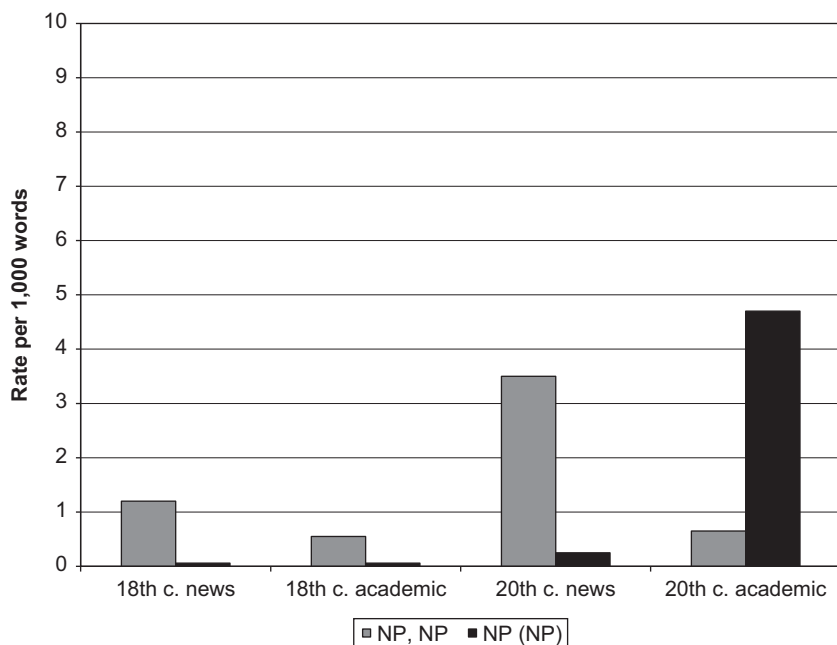


Figure 4.10 Appositive noun phrases as nominal post-modifiers

*In about two months after it had acquired this additional head, a fragment separated from the tail (**the most usual place of separation**) and was in progress towards its entire reproduction when it was accidentally lost . . .*

*the former is composed of sporules, empty tubes (**the mycelium**), and tubes filled with sporules*

As Figure 4.10 shows, this textual device has become quite common in the twentieth century, but it is restricted almost entirely to science research writing.

In this section, we have considered the magnitude of historical change in the use of the phrasal devices that function as noun modifiers, by examining the overall frequency of those features in academic science writing. This analysis provides a complementary perspective to the correlational analyses presented in Sections 4.3.1–4.3.3 (which measure the extent to which change has been systematic). The analyses in the present section have shown that most of these phrasal noun modifiers are extremely frequent in academic texts, and that they have increased in use dramatically during the twentieth century.

4.3 What features have decreased in use in academic prose?

Given the strong historical increase in the use of nouns and phrasal noun-modifying structures in academic prose, it is natural to wonder what features have decreased in use. The quantitative findings presented in [Figures 4.3–4.10](#) are rates of occurrence per 1,000 words of text. So, if one feature increases in use, other features must decrease in use by the same amount. For example, [Figure 4.4](#) shows that nouns in science writing increased in use during the twentieth century by about 125 per 1,000 words. Thus, if we were to compare 1,000-word samples from different historical periods, we would find the following: in a 1900 science research article, 265 of those words were nouns, while the rate for nouns rose to almost 400 of the 1,000 word sample in research articles published in 2000. These rates can be interpreted in a mechanical way: In a typical nineteenth century science text, c. 25% of the words were nouns, while in a typical twenty-first century science text, c. 40% of the words were nouns. Thus, an additional 15% of a typical present-day science text is realized as nouns.

As a result, other features must be used less commonly, to make room for the additional nouns. One obvious choice might be verbs: as the density of nouns goes up, we might expect the density of verbs to go down. [Figure 4.4](#) shows that the direction of change for lexical verbs in science prose is as predicted, but not the magnitude of that change. That is, lexical verbs have decreased in use by only c. 20 per 1,000 words (or only by c. 2% of the text).

However, [Tables 4.2 and 4.3](#) show that there have been several other grammatical features that have decreased in use in science prose, including adverbs, pronouns, and prepositions. It turns out that some of these changes have been relatively large in magnitude. For example, [Figure 4.8](#) shows that adverbial prepositional phrases have decreased by almost 30 per 1,000 words over the last century, and *of*-phrases have decreased by almost 20 per 1,000 words. [Table 4.3](#) shows a large negative correlation for third person pronouns in academic prose. That correlation corresponds to a dramatic decline in the frequency of third person pronouns, from a rate of c. 20 per 1,000 words in the eighteenth century to a miniscule rate of only 2–3 per 1,000 words in the late twentieth century.

In addition, other features have dramatically declined in use but escaped the notice of previous analysts. One of the most surprising changes in academic discourse style is a strong decrease in the use of definite articles (and determiners generally). We expected that determiners would increase in use, accompanying the strong increase in nouns in science writing. However, the actual historical development has been exactly the opposite,

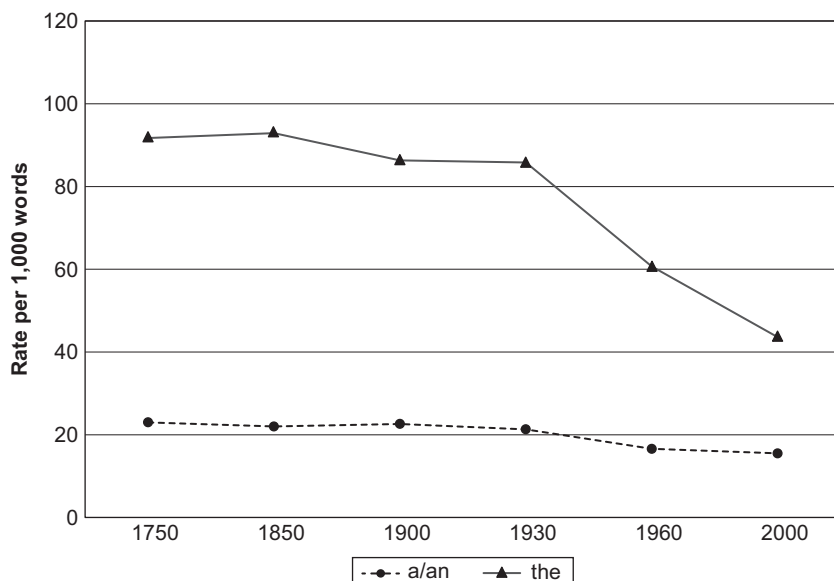


Figure 4.11 Historical change in the use of definite and indefinite articles in science academic prose

as shown in Figure 4.11. Indefinite article use has been declining gradually over the past century. But the dramatic shift here is for definite articles, which were slowly declining over the course of the nineteenth century, and then declined dramatically in the second half of the twentieth century.

This reduction in definite article use has several sources, including a greater use of abstract process nouns and technical nouns (with a function similar to proper nouns, which typically occur without determiners). The use of abstract/process or technical head nouns without determiners was already prevalent in nineteenth century science writing, as in Text Sample 4.3. For example, nouns like *neutralization* and *dyeing* refer to processes and thus do not normally take a determiner, while nouns like *hydrate*, *alumina*, *magnesia*, *carbonate*, and *lead* are essentially proper nouns.

Text Sample 4.3

Nineteenth century science research writing

Definite articles are marked in *italics*; noun phrases with no determiner are **bold underlined**.

By adding a variety of **substances** to an extract of **madder** with **cold water**, I was enabled to ascertain under what circumstances and by

what means *the* tinctorial power of *the* liquid is destroyed, and consequently what is *the* general character of *the* substance or substances to which it is due. I found that by adding sulphuric or muriatic acid to the extract and heating, *the* liquid, after **neutralization** of *the* acid was no longer capable of **dyeing**. *The* tinctorial power was also destroyed by *the* addition of hydrate of alumina, magnesia, protoxide of tin and various metallic oxides, but not by carbonate of lime or carbonate of lead. In all cases in which *the* property of **dyeing** in *the* extract was destroyed, I invariably found that its bitter taste and bright yellow colour were lost.

Edward Schunck. 1851.

'On Rubian and Its Products of Decomposition'

Philosophical Transactions, 141, 433–459.

These same types of nouns occur even more frequently without determiners in present-day academic prose. For example, [Text Sample 4.4](#) illustrates the dense use of process nouns (e.g., *selection*, *reproduction*, *growth*) and other abstract nouns without determiners (e.g., *synchrony*, *variation*, *phenology*, *predictability*, *constancy*).

Text Sample 4.4

Twentieth century science research writing

Definite articles are marked in *italics*; noun phrases with no determiner are **bold underlined**.

Summary

1. **Timing** and **synchrony** of **reproduction** are regarded as crucially important **factors** for **fitness** in seasonal **environments**. Natural **selection** has probably favoured temperate and arctic **female herbivores** that match **reproduction** with **onset** of **plant growth** in **spring**. However, **breeding synchrony** may also be affected by **variation** in phenotypic **quality** of **females** in a population, because **females** in poor **body condition** have been found to delay **ovulation** and subsequent **calving**.
2. We compared **breeding phenology**, i.e. *the* timing and synchrony of **rutting** (**roaring**, sexual **aggregation**) and **calving** of red **deer** (*Cervus elaphus* L.) in **France** (**latitude**: 49°N) and **Norway** (**latitude**: 63°N). We hypothesized (H1) that **calving** and **rutting** were later at the site with latest **onset** of **plant growth**.
3. We further quantified overall environmental **predictability** as the sum of annual **constancy** and **seasonality** and tested three different (not mutually exclusive) hypotheses about **breeding synchrony**: (H2a) *the* population experiencing most seasonal **plant phenology** should show *the* highest breeding synchrony; (H2b) overall **predictability** of **plant phenology** should determine **breeding**

synchrony; and (H2c) **breeding** should be more synchronized in *the* population with lowest **female body weight variation** within **age classes** because they ovulate more synchronously.

L.E. Loe et al. 2005.

‘Climate Predictability and Breeding Phenology in Red Deer: Timing and Synchrony of Rutting and Calving in Norway and France’
Journal of Animal Ecology, 74: 579–588.

In addition, there has been a shift to omit definite articles even for noun phrases that can be regarded as specific in reference, apparently to generalize the reference to a whole population rather than the specific objects considered in a particular study. Such factors can help to explain the absence of the definite article in phrases like:

the site with Ø latest onset of plant growth

the population experiencing Ø most seasonal plant phenology

Ø overall predictability of plant phenology should determine breeding synchrony

the population with Ø lowest female body weight variation within age classes

Thus, the increased use of nominal/phrasal structures in academic writing has been offset by decreases in a range of other devices. These include features that are clearly complementary to noun phrase structures, such as lexical verbs, auxiliary verbs, pronouns, and adverbs. But more surprisingly, the features that have declined in use also include prepositions (especially *of* and also other prepositions functioning as heads of adverbial phrases) as well as determiners. We return to more detailed discussions of these developments in [Chapters 5](#) and [6](#).

4.4 Historical change across sub-registers within academic research writing

One primary goal in this book is to highlight the importance of academic prose for historical linguistic research, because it has been so innovative in the development and use of phrasal modifiers. These changes are much more pronounced in academic writing than in most other registers of English, directly contradicting the stereotypical view that academic writing is resistant to grammatical change.

The changes in discourse style documented in the preceding sections have been rapid historical developments, occurring mostly in the twentieth century. However, as we showed in [Chapter 3](#), different sub-registers within present-day academic writing differ in the extent to which they

exhibit these characteristics: academic research articles employ greater structural compression than university textbooks; science research articles employ greater structural compression than humanities research writing; specialist science research articles employ greater structural compression than popular science articles.

In the present section, we track the historical evolution of such differences, showing how sub-registers of academic prose have become increasingly diversified over the course of the twentieth century. In particular, we compare the historical development of four specific academic sub-registers: specialist science research articles, popular multi-disciplinary science research articles, social science research articles, and humanities articles/books.

In the eighteenth and nineteenth centuries, these differences were less important: social science was not well-established as a separate discipline; there were relatively few well-established sub-disciplines of science; and most science writing was directed towards a multi-disciplinary readership (see discussion in the following paragraphs). Thus, for our analyses in this section, we distinguish between only two general nineteenth century sub-registers of academic prose: multi-disciplinary science research articles and humanities books (history). In contrast, disciplinary differences become much more important in the twentieth century. As a result, we include a more specific sampling of academic sub-registers for the twentieth century, with texts from specialist science research articles, popular multi-disciplinary science research articles, social science research articles, and humanities (history) articles/books. A full description of these sub-corpora is provided in [Chapter 2, Section 2.3](#).

The comparison of specialist versus generalist multi-disciplinary science writing requires some background discussion. Some of the science research journals that were influential in the eighteenth and nineteenth centuries have continued to publish research articles up to the present day. For example, the *Philosophical Transactions of the Royal Society (PT)* has been published continuously since 1665, and the journal *Science* has published research articles since 1880. As a result, previous studies have tracked linguistic change in English science writing through consideration of these influential research journals (see, e.g., Biber, Finegan, and Atkinson 1994; Atkinson 1996; Biber and Finegan 1997/2001; Valle 1999; Gross et al. 2002). In many ways, these journals have remained relatively constant in purpose and readership, intended as outlets for the most important science research findings from across the full range of scientific disciplines, and written for an audience from that same breadth of disciplines. By basing historical analyses on a single research journal, which has remained

relatively constant in purpose and readership across time, researchers have been able to isolate the influence of historical change (i.e., minimizing any confounding influence that would result from comparisons of different academic journals).

However, an additional consideration is the role of the targeted research journal relative to the universe of academic research writing in a given historical period. Thus, in the eighteenth and nineteenth centuries, most science writing was published in a relatively small number of multi-disciplinary journals, and read by a wide multi-disciplinary audience. Research articles published in journals like *PT* and *Science* provide a fairly good representation of the universe of science writing from those periods. In contrast, there has been an incredible proliferation of academic research journals in the last half of the twentieth century, and as a result, generalist journals like *PT* and *Science* represent only a very small sector of present-day science writing. Most science writing today is published in highly specialized journals associated with specific sub-disciplines, and read mostly by specialists from those same sub-disciplines. As we show next, this difference actually has important linguistic consequences: science research articles published in specialist research journals from a specific sub-discipline turn out to be quite different grammatically from articles published for a more general readership in multi-disciplinary journals like *PT* and *Science*. And as a result, the dramatic historical changes towards increased use of compressed phrasal devices is considerably more pronounced in specialist science research writing than in any other register.

As the preceding analyses have shown, one of the most salient grammatical characteristics of modern academic writing is the heavy reliance on nouns. Figure 4.3 shows that the density of common nouns has been increasing steadily in academic writing over the past 200 years, with a more rapid increase during the twentieth century. This shift towards a more “nominal” style was noticed as early as 1960 by Wells. However, the fact that this is an ongoing change, which has intensified in the last half of the twentieth century, means that Wells’s observations in the 1960s represent a density of nominal use that is considerably less extreme than in present-day academic research writing!

Figure 4.12 shows that academic sub-registers have participated in this change to differing extents. In particular, history research writing has changed little over this period: it had a relatively dense use of nouns in the nineteenth century, and it has maintained roughly that same density of nouns up to the present time. In contrast, the use of nouns has increased dramatically in science research writing, and it appears that the increase is still in progress. The trend is by far most evident in

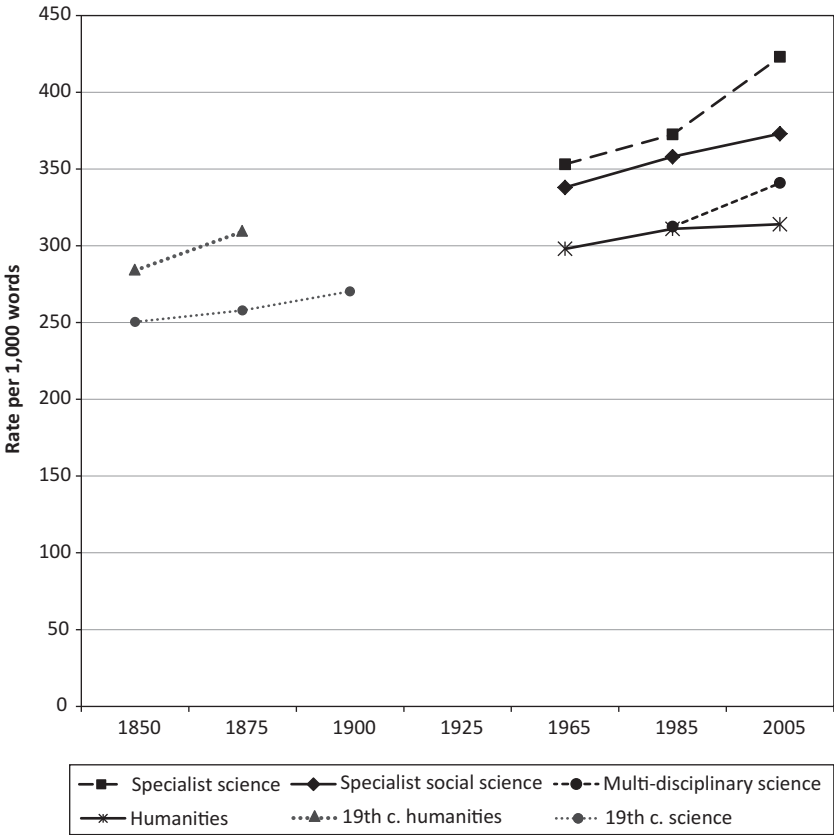


Figure 4.12 Historical change in the use of nouns across academic sub-registers

specialist science research articles, where nouns have increased by over 10% in just the last 20 years. Specialist social science articles follow the same increasing trend as specialist science articles, but to a less extreme extent. The trend towards increased use of nouns has also affected multi-disciplinary science articles, but it is considerably less marked than in either of the specialist sub-registers.

Appendix 2 presents the actual mean scores and standard deviations for nouns (and several other grammatical features) across academic sub-registers in recent historical periods. The first row in Table 4.4 presents the results of a factorial ANOVA to test the statistical significance of the mean differences for nouns across three recent historical periods (1965, 1985, 2005) and across the four academic sub-registers. Both main effects

Table 4.4 *Summary of the ANOVA factorial models for four academic sub-registers (specialist science, specialist social science, multi-disciplinary science, humanities) in three recent historical periods (1965, 1985, 2005)*

	Model F-Score	Model Significance	Model R ²	Year	Sub- Register	Year * Sub- Register
Nouns	51.0	<.0001	0.477	<.0001	<.0001	<.0001
Nominalizations	7.9	<.0001	0.124	<.0001	<.0001	ns
Relative Clauses	33.6	<.0001	0.376	ns	<.0001	ns
Noun + <i>of</i> -phrase	10.5	<.0001	0.158	<.0001	<.0001	ns
Noun + Noun	71.1	<.0001	0.560	<.0001	<.0001	<.001

(historical period and sub-register) show significant differences. In addition, there is a significant interaction effect here, reflecting the fact that humanities research writing has changed little in the use of nouns, while the frequency of nouns has increased dramatically in the science sub-registers (see Figure 4.12).

Nominalizations have followed a quite different developmental path from common nouns. Figure 4.13 shows that the frequency of nominalizations has increased in all sub-registers of academic writing. Surprisingly, though, the pattern of sub-register diversification is the opposite of that found for common nouns: nominalizations have increased the most in specialist social science articles and in multi-disciplinary science articles, and they have also increased considerably in humanities research writing. In contrast, the increase is considerably less pronounced in specialist science articles. (Table 4.4 shows that these mean differences are statistically significant for both historical period and sub-register.)

The discussion in preceding sections has documented a general decrease in the use of clausal modifiers, complementing the increase in nominal/phrasal features. For example, Figure 4.7 shows that relative clauses have decreased in both fiction and academic prose over the last century. Figure 4.14 provides additional information about this trend, showing that there have been distinct patterns of change among academic sub-registers: The decrease in use is by far strongest in specialist science articles, where relative clauses have declined by 50% from 1900 to 1965. Specialist social science articles also show a strong decline in use, while the multi-disciplinary science articles show a lesser decline in use. And in contrast to all science sub-registers, history

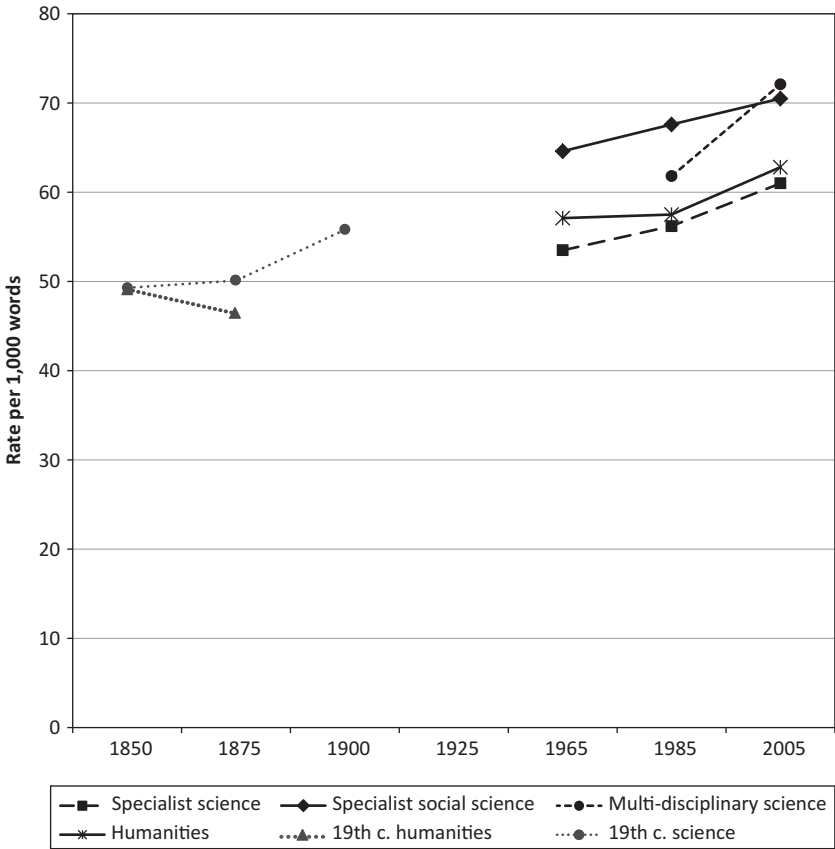


Figure 4.13 Historical change in the use of nominalizations across academic sub-registers

research writing has remained essentially unchanged over the past century in its heavy reliance on relative clauses for nominal elaboration. As a result, relative clauses are at present considerably more common in humanities academic prose than in any of the science registers.²

Figure 4.15 shows that post-nominal *of*-phrases have followed a similar historical path: strong declines in specialist science and social-science

² The ANOVA results reported in Table 4.4 are based on historical change for the three recent historical periods – 1965, 1985, and 2005. Thus, the results for relative clauses show a significant difference for sub-register, but no significant difference for historical period. That is, the large historical decrease in the use of relative clauses in science writing occurred in the early twentieth century, while these registers have remained relatively unchanged in their infrequent use of relative clauses over the past fifty years.

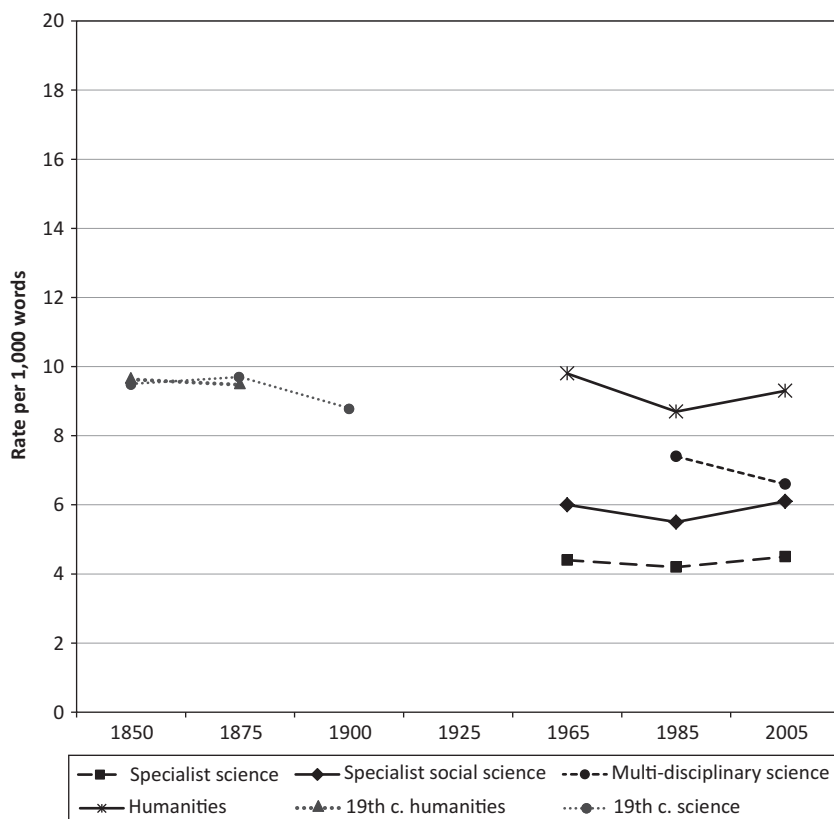


Figure 4.14 Historical change in the use of relative clauses across academic sub-registers

articles; a lesser decline in multi-disciplinary science articles; and essentially no change in history research writing. Those trends are in marked contrast to historical change in the use of post-nominal phrases headed by other prepositions (e.g., *in*, *on*, *for*, *with*), which have increased dramatically in science research writing over the last century (see Figure 4.9; cf. the discussion of Figure 3.12 in Chapter 3).

These findings (together with the findings in Section 3.3.4) show that there are important linguistic differences among academic sub-registers, and thus the cover term ‘academic prose’ is not so useful in referring to a grammatical discourse style. For example, the discourse of humanities research writing has changed little over the past three centuries: it still

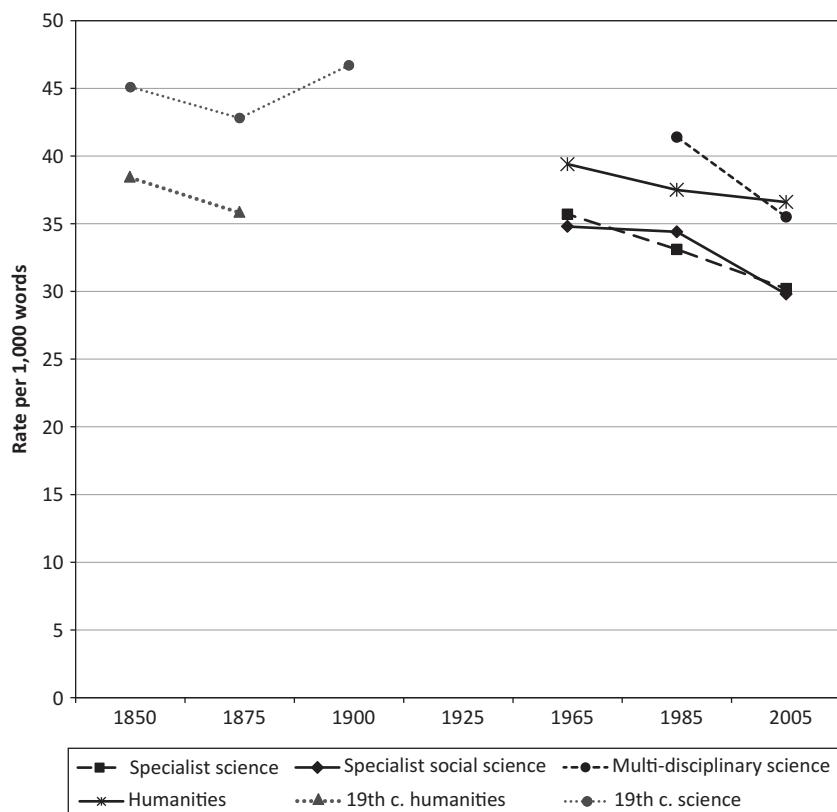


Figure 4.15 Historical change in the use of noun + *of*-phrase across academic sub-registers

employs frequent dependent clauses, and makes comparatively little use of phrasal noun modifiers. In contrast, the discourse of science research writing has undergone extensive change, shifting from a clausally-elaborated style to a style employing a dense use of noun phrase modification.

In addition, the most important linguistic changes have not been associated with the grammatical features usually targeted in previous research. Two of the most salient grammatical devices used for noun phrase modification (relative clauses and *of*-genitives) have actually declined in use in specialist science writing. Nominalizations have increased only slightly in specialist science research articles, while they have increased to a greater extent in social science and multi-disciplinary science articles.

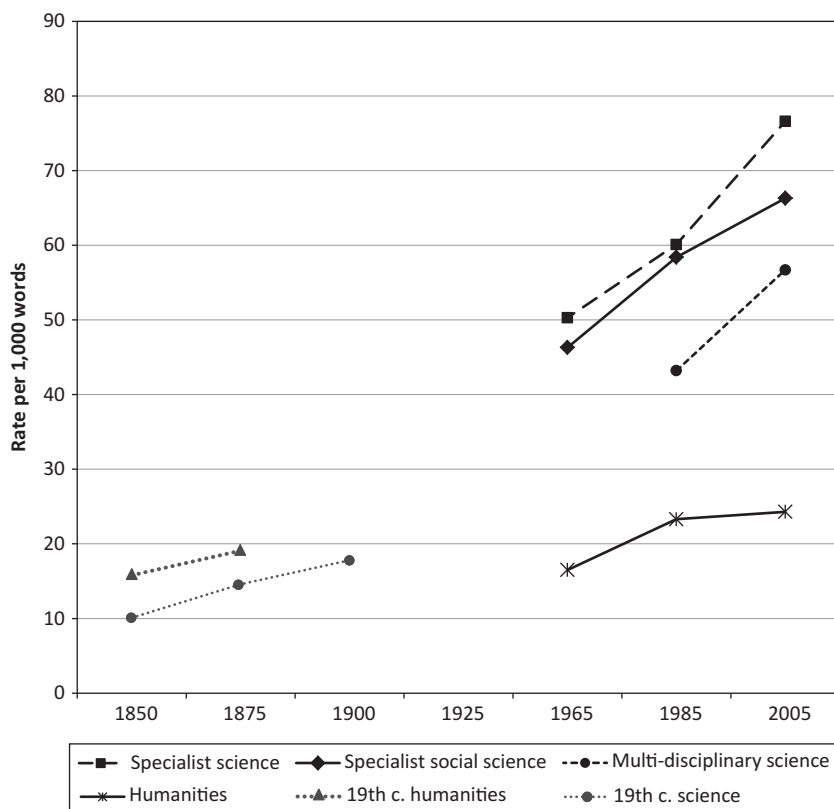


Figure 4.16 Historical change in the use of nouns as nominal pre-modifiers across academic sub-registers

So what has increased in use in specialist science research writing, resulting in the distinctive discourse style documented in [Chapter 3](#)? Probably the most important grammatical development has been the dramatic increase in the use of nouns used as a pre-modifier of a head noun (e.g., *patient history*, *case study*). As [Figure 4.16](#) shows, pre-modifying nouns were only moderately common in the nineteenth century, but they have increased in use by as much as 400% in the twentieth century. This increase has been strongest in specialist science research writing, and also very strong in specialist social-science writing. Multi-disciplinary science writing is progressing in the same direction, but it lags behind the specialist sub-registers in the extent to which this device is employed. In contrast, history research writing has increased only slightly in the use of this

feature. These developments are significant across historical periods and across sub-registers (see [Table 4.4](#)). In addition, there is a significant interaction effect for this feature, reflecting the fact that humanities writing has changed little, in contrast to the dramatic increase over time observed for science writing.

4.5 Chapter summary

In summary, there are important differences across sub-registers of academic research writing in the extent to which they have permitted grammatical innovations associated with the increased use of phrasal devices modifying head nouns. In particular, we have shown that specialist science research articles have been the most ‘agile’ academic register (to employ the term used by Hundt and Mair [1999](#)), exhibiting a remarkable change in discourse style over the course of the twentieth century. Thus, far from being ‘up-tight’ and resistant to change, we have shown how specialist science research articles have changed dramatically in their use of phrasal grammatical devices associated with informational ‘compression’. It turns out that these are not merely changes in frequency of use: in [Chapter 5](#), we discuss the functional extensions associated with these quantitative trends.

The functional extension of phrasal grammatical features in academic writing

5.1 Introduction

The research findings presented in the last two chapters have been primarily quantitative in nature, describing the extent to which academic research writing employs phrasal grammatical devices, in contrast to the grammatical discourse styles of other registers. In [Chapter 3](#), we began with the fundamental opposition between clausal discourse styles and phrasal discourse styles, showing that conversational discourse relies mostly on clausal grammatical features (including a dense use of dependent clause features), while academic writing relies primarily on phrasal grammatical features (especially those functioning as noun phrase modifiers). This opposition is best described in terms of a continuum of use influenced by three major situational factors: production in the spoken versus written mode, communicative purpose (especially personal/interpersonal versus informational), and the extent to which discourse is written by specialists and addressed to other specialists. The registers that are the most marked situationally are also the most distinctive in their linguistic discourse styles. Thus, conversation is produced in the spoken mode; it has personal/interpersonal communicative purposes; and it is a register that all speakers of a language participate in. Linguistically, conversation is distinctive in its frequent use of ‘colloquial’ features (e.g., contractions, pronouns, discourse markers) and also in its frequent use of finite dependent clauses functioning as clause constituents. At the other extreme, academic research writing is produced in the written mode; it has informational communicative purposes; and it is a register written by specialists and intended for other specialists. Linguistically, academic prose has the opposite characteristics from conversation: very few ‘colloquial’ features, and few finite dependent clauses (especially those functioning as clause constituents), but a frequent use of phrasal features functioning as noun phrase modifiers. There is a large set of related phrasal devices that pattern in this way, including nouns

as nominal pre-modifiers, attributive adjectives, noun–participle compounds as nominal pre-modifiers, prepositional phrases as nominal post-modifiers, and appositive noun phrases as nominal post-modifiers.

Chapter 4 added further information about these register differences, showing that the phrasal grammatical style typical of present-day academic research writing is actually a recent innovation. In fact, this grammatical style is not attested in any register in the eighteenth and even early nineteenth centuries, and thus it represents one of the most important grammatical changes in English of the last four centuries. However, conforming to the perspective described by Fennell (2001) and Denison (1998), the historical changes documented in Chapter 4 are quantitative/statistical trends rather than completely new grammatical innovations. That is, those historical developments involve structures that have become ‘more or less common generally or in particular registers’ (Denison 1998: 93). Thus, phrasal devices functioning as noun phrase modifiers were well-attested in the seventeenth and eighteenth centuries, but they were generally rare at that time in all registers. Over the following centuries, they have gradually increased in use in informational written registers, while their frequency of use remained virtually unchanged in other registers. Then, in the twentieth century, we witnessed a dramatic increase in use, restricted primarily to informational written registers.

It is clear from the quantitative findings presented in Chapters 3 and 4 that these grammatical trends have resulted in a new discourse style not attested in earlier historical periods. This grammatical style is innovative in the ways in which it packages information in discourse: (1) presenting information in noun phrases rather than in clauses, and (2) presenting information in phrasal modifiers rather than in clausal modifiers.

Most registers in English employ clausal discourse styles, and historically, clausal discourse styles have been the norm. In fact, some kinds of academic prose continue to employ clausal styles even up to the present time. For example, humanities research writing, with its narrative and descriptive communicative purposes, makes extensive use of a discourse style with many clausal modifiers that function as adverbials or objects of verbs. Finite relative clauses are also extremely frequent in this sub-register; these are often non-restrictive, providing extra information rather than identifying the reference of the head noun. The following example sentences from history research writing illustrate this type of complex academic prose with multiple levels of embedded dependent clauses:

The French **tended** [to **trap** and **trade** for furs], but the British settlers, [who were starting [to survey land in [what is now Tennessee and Kentucky]], looked like incipient farmers.

[As Washington **noted** in his journal], the French army officers, [after they **had** "**dosed** themselves pretty plentifully with wine," [which "**gave** a License to their Tongues [to **reveal** their Sentiments more freely,"]]] **told** Washington: ["It **was** their absolute Design [to **take** possession of the Ohio], & by G– they **would do** it."]

In contrast, science (and social science) research writing tends to employ phrases rather than clauses, with information packaged as noun modifiers rather than as clause constituents. The following sentence from a psychology research article is typical:

This may indeed be part of the reason for the statistical link between schizophrenia and membership in the lower socioeconomic classes.

The clausal syntactic structure of this sentence is extremely simple, with only one main verb phrase:

X **may be** Y (*This may be part*)

Instead of embedded clauses, all of the information is packaged as phrasal modifiers embedded in noun phrases:

*This may indeed be **part** [**of** the **reason** [**for** the statistical **link** [**between** schizophrenia and **membership** [**in** the lower socioeconomic classes]]]]*.

Our discussion up to this point has described the historical development of this new grammatical discourse style from a quantitative perspective. In the present chapter, we take the analysis of these features one step further, showing that these changes represent more than just a shift in preferred discourse style. That is, contrary to the expectations raised by Fennell (2001) and Denison (1998), we argue here that these historical developments are *not* merely quantitative/stylistic trends. Rather, there have been important extensions to the grammatical/discourse functions of these devices accompanying the increases in frequency of use. These functional extensions can be documented for each of the individual grammatical devices described in the preceding chapters. In Section 5.2, we explore the use of phrasal grammatical devices that function as nominal pre-modifiers, and then in Section 5.3, we turn to a detailed consideration of devices functioning as nominal post-modifiers.

5.2 Phrasal features functioning as pre-modifiers of a head noun

5.2.1 *Nouns as nominal pre-modifiers*

The shift towards the present-day reliance on nouns as nominal pre-modifiers in academic written prose is perhaps the most dramatic quantitative–historical change in English witnessed in the last three centuries. As Figures 4.5 and 4.16 showed, pre-modifying nouns were rare in the seventeenth, eighteenth, and even nineteenth centuries, but then they rapidly increased in use over the course of the twentieth century, to the extent that they are now ubiquitous in modern written academic texts.

This change is not merely a quantitative shift in stylistic preference. From a discourse perspective, texts are now constructed in fundamentally different ways from earlier centuries, relying to a large extent on phrasal devices rather than dependent clause structures. But deeper consideration reveals linguistic change at a more fundamental level, involving a steady expansion of meaning and function. In the present section, we describe five factors¹ related to the functional extension of pre-modifying nouns:

- A) pre-modifying nouns used as an alternative to *'s*-genitive and *of*-genitive phrases
- B) semantic classes of the nouns used as pre-modifiers
- C) nominalizations as pre-modifiers
- D) sequences of multiple pre-modifying nouns
- E) the meaning relationships between nouns

To investigate these factors, we undertook detailed qualitative analyses of noun–noun sequences in science prose, medical prose, and newspaper articles from the ARCHER Corpus.

¹ Noun–noun sequences can additionally be analyzed for a number of syntactic distinctions. For example, Huddleston and Pullum (2002: 448ff) distinguish between NN sequences functioning as 'composite nominals' (e.g., *lemon sorbet*, *microfilm reader*) versus sequences that represent compounds (e.g., *ice-cream*). This distinction is related to the status of the pre-modifying noun as an 'adjunct' (or 'modifier') versus 'complement' (e.g., *London newspapers* vs. *television screen*); see e.g., Rosenbach (2007: 146–147); Huddleston and Pullum (2002: 449). Rosenbach (2007) further distinguishes between 'classifier' versus 'determiner' functions for pre-modifying nouns that are complements. While it is possible to cite clear-cut examples for these distinctions, many natural occurrences of NN sequences are intermediate and difficult to classify. Thus, consideration of such syntactic factors is beyond the scope of the present study.

A) *Pre-modifying nouns used as an alternative to 's-genitive and of-genitive phrases*

One major reflection of the functional extension of pre-modifying nouns is that they are now commonly used to express genitive meanings, replacing the functional domain of *'s*-genitive and *of*-genitive phrases. Thus, consider the following example from a newspaper article:

... *the Pope met [...] the Communist Party chief...*

The final noun phrase in this example expresses a genitive relationship that could be paraphrased with either of the other two genitive variants:

noun-'s + head noun: *the Communist Party's chief*

head noun + *of*-phrase: *the chief of the Communist Party*

Previous studies have argued that *'s*-genitives have been historically expanding in use at the expense of the *of*-genitive (see, e.g., Rosenbach 2002; Hinrichs and Szmrecsanyi 2007; Leech et al. 2009: chapter 10.4). These studies, which are based on analysis of linguistic tokens that are interchangeable, show strong patterns of historical change during the twentieth century. For example, Hinrichs and Szmrecsanyi (2007: 448) show that in 1961, *'s*-genitives accounted for only c. 37% of all interchangeable *'s*-genitives and *of*-genitives in the newspaper registers from the Brown and LOB corpora. By 1991, that pattern of use had changed so that *'s*-genitives accounted for c. 50% of all interchangeable *'s*-genitives and *of*-genitives in newspaper texts from the Frown and FLOB corpora.

These previous studies have been restricted to only two linguistic variants: *'s*-genitives versus *of*-genitives. However, when pre-modifying nouns are considered as a third genitive option, a different picture emerges (see Biber et al. 2016; Szmrecsanyi et al. 2015), with pre-modifying nouns becoming the preferred option for expressing genitive relationships.

The findings presented in Chapters 3 and 4 are based on the textual rates of occurrence for grammatical features (i.e., rather than the percentage use of a variant as a proportion of interchangeable occurrences). This type of research design has been called a 'text-linguistic' perspective on linguistic variation, because it tells us how often we will encounter a linguistic feature in texts (see Biber 2012; Biber et al. 2016). Text-linguistic findings regarding post-nominal *of*-phrases show that this feature has declined in use in fiction and newspaper prose (see Table 4.3). Although the historical change for *of*-phrases is less strong in the sub-registers of academic prose, there has been a consistent historical decline in those registers as well, with the strongest declines seen in science and social science research articles (see Figures 4.8

and 4.15). In contrast, pre-modifying nouns have increased dramatically in use from a text-linguistic perspective, and that increase has been strongest in science research articles (see Table 4.3; Figures 4.5, 4.16). The text-linguistic rates of occurrence for *'s*-genitives are not reported in Chapters 3 and 4, because they are quite rare in academic writing (only c. 2.5 per 1,000 words in present-day academic prose; see Biber et al. 1999: 302).

The text-linguistic perspective can be contrasted with a variationist perspective, based on analysis of linguistic tokens that are interchangeable with other structural options (see discussion in Biber et al. 2016; Szmrecsanyi et al. 2015). Variationist research designs are based on analyses of each token of a linguistic feature, identifying the particular linguistic variant used in that occurrence, and predicting the use of linguistic variants by reference to other contextual factors. The results in such analyses are proportional, comparing the percentages of a particular variant out of all interchangeable constructions.

From a variationist perspective, *'s*-genitives are expanding in use relative to *of*-genitives (see earlier discussion). However, when we include pre-modifying nouns as a third variant, it turns out that this newer structural variant is expanding in use relative to both *'s*-genitives and *of*-genitives.

Thus, Figure 5.1 plots the proportional use of interchangeable nouns as pre-modifiers (versus *'s*-genitives). In most cases, these two construction types are in general not interchangeable. As a result, the proportions shown in Figure 5.1 are based on very small samples, causing some fluctuations in the historical trends. For example, the nineteenth-century proportions of 55% pre-modifying nouns (versus 45% *s*-genitives) in science prose is based on a sample of only forty tokens that occurred in the ARCHER Corpus (22 noun pre-modifiers that are interchangeable with *s*-genitives, and 18 *s*-genitives that are interchangeable with noun-pre-modifiers). However, despite the relatively small sample sizes, the overall historical trends are consistent across registers, with a notable increase in the proportional use of noun pre-modifiers (and thus a corresponding decline in the proportional use of *'s*-genitives) across the centuries. Science prose shows the strongest increase, with noun pre-modifiers being used over 95% of the time in interchangeable constructions in the twentieth century.

The sample of interchangeable occurrences in ARCHER for the *of*-genitive versus noun pre-modifier alternation is much larger (see Biber et al. 2016; tables 2 and 3), and correspondingly, the historical trends shown in Figure 5.2 are much more consistent across centuries. For all three registers, there is a strong increase in the proportional use of noun pre-modifiers at the expense of *of*-genitives. Letters and newspaper

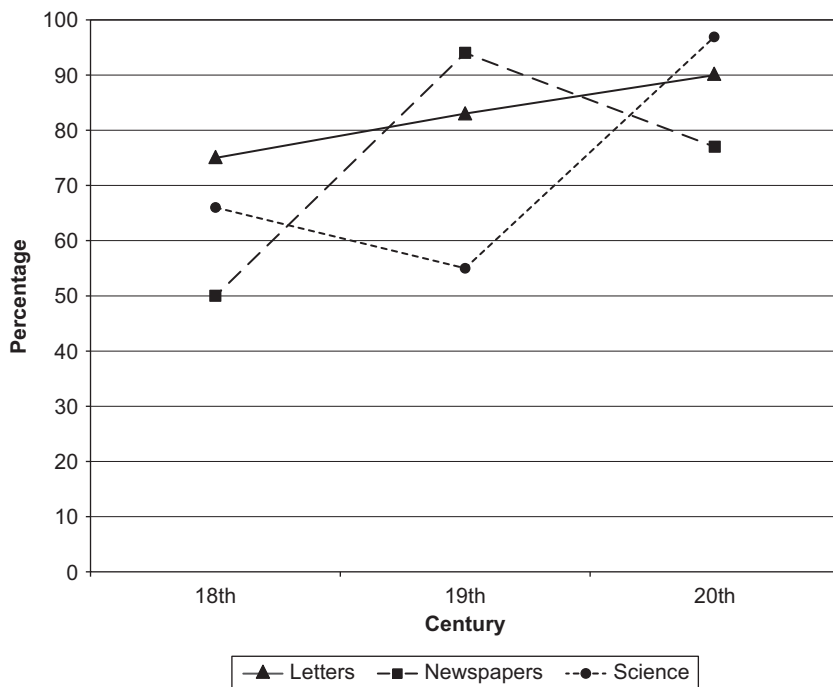


Figure 5.1 Historical change in the proportional use of pre-modifying noun genitives (versus 's-genitives)

Based on all occurrences of pre-modifying nouns and 's-genitives that are interchangeable

reportage take the lead in this change during the nineteenth century, while science prose shifted strongly towards increased noun pre-modifier variants during the twentieth century.

These findings provide strong evidence for the functional expansion of pre-modifying nouns, related specifically to the increasing use of these features to express genitive meanings. In [Chapter 4](#), we discussed the increased use of 'other' prepositional phrases as nominal post-modifiers (e.g., *in*, *on*, *for*, *with*), and compared that trend to the overall decline in frequency of *of*-phrases as nominal post-modifiers (see detailed discussion in [Section 5.3.1](#)). In the present section, we have shown that this decline in use is also related to the historical increase of nouns used as pre-modifiers, related in part to the functional extension of these structures to express the same kinds of genitive meanings that have traditionally been expressed by *of*-phrases (and the 's-genitive). We return to this wider perspective (i.e., considering both pre-modifying and post-modifying structures) in [Section 5.4](#).

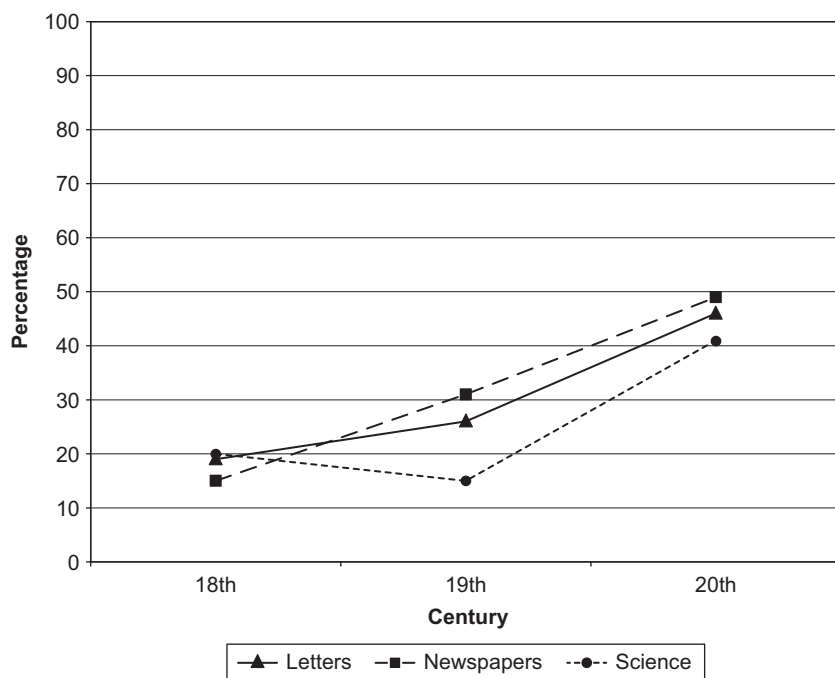


Figure 5.2 Historical change in the proportional use of pre-modifying noun genitives (versus *of*-genitives)

Based on all occurrences of pre-modifying nouns and *of*-genitives that are interchangeable

B) *Semantic classes of the nouns used as pre-modifiers*

In the sixteenth century, the functions of pre-modifying nouns were extremely restricted, with most occurrences referring to titles such as *King*, *Master*, or *Doctor* (see Raumolin-Brunberg's 1991 analysis of noun phrase structures in Sir Thomas More's writings). In the seventeenth and eighteenth centuries, pre-modifying nouns came to be used for an expanded set of functions, but they were still quite restricted in use, expressing meanings related to three general semantic categories: title nouns, place/location nouns, and concrete/tangible nouns. Following are typical examples from seventeenth and eighteenth century written informational prose.

Title nouns:

King, Doctor, Duke, Captain, Lord, and so on

Place nouns:

Specific places:

Hampton Court, Dumbarton Castle, India Company, Greenwich Park,
Boston papers, London gazette

General locations:

country habitations, farm houses, field marshall, frontier garrisons,
ground floors, home affairs, town wall

Concrete/tangible nouns:

cannon ball, castle hill, chamomile tea, coffee house, copper mill, corn field,
foot soldiers, flannel roller, fountain water, goose eggs, gun ship, hen eggs,
iron chain/particles/tools, linen handkerchief, milk diet, mineral taste/
spring(s)/poisons, sand bank, sea captains, spring water, tea commissioners

By the late-1800s, when the overall frequency of noun–noun sequences began to increase, there was also a marked expansion in the range of meanings expressed by pre-modifying nouns. Thus, in addition to the categories described earlier, we find pre-modifying nouns commonly used to refer to institutions, states or conditions, and other intangibles.

Institutions:

family history, school proposal, state convention, union member

States or conditions (often diseases in medical prose):

cancer cells, croup cases, diphtheria results, health department, maternity hospitals, smallpox eruption

Other intangibles:

class examinations, currency troubles, credit foncier, heat apoplexy,
weather bureau, temperature chart, quarantine restrictions

This expansion in the kinds of meanings expressed by pre-modifying nouns continued over recent decades, so that by the end of the twentieth century, almost any noun could be used freely as a pre-modifier of another head noun. These include a wide range of abstract nouns with intangible meanings, such as:

age group, casualty department, emergency powers, income tax, monopoly act, news agency, peace conference, press conference, price commission,
sector strike, sex differences, television interview, time interval, wage increases,
weight loss

C) *Nominalizations as pre-modifiers*

A related historical change – beginning in the middle of the nineteenth century – involves the use of nominalizations in noun–noun sequences, both as the head noun (e.g., *school proposal*, *class examinations*, *quarantine restrictions*) and as the pre-modifying noun. For the most part, pre-modifying nominalizations refer to processes or activities; they are derived from verbs, either through morphological derivation or through conversion. This expansion in use continued throughout the twentieth century, with a much greater range of nominalized and converted forms derived from verbs being used as noun pre-modifiers. Consider the following examples.

Twentieth century nominalizations and noun conversions as NP pre-modifier, referring to processes or activities:

Nouns morphologically derived from verbs:

extradition treaty, government officials, inoculation experiments, insurance companies, investigation department, publication house, service reform, taxation prospects, correlation coefficients, population base, regression analysis, reprisal raids, terrorism centre

Verb to noun conversion:

awards bureau, murder trials, research fund, trade agreement, transport unions, study period

In addition, nominalized forms derived from nouns or adjectives also came to be commonly used as noun pre-modifiers in the twentieth century. These nominalizations typically refer to abstract attributes or qualities rather than processes. For example, consider the following:

Nominalizations referring to abstract attributes or qualities:

freedom movement, intelligence agencies, majority group, memorial service, mortality rate, safety officials, security interests

D) *Sequences of multiple pre-modifying nouns*

A fourth type of functional extension is the increased use of noun phrases with multiple pre-modifying nouns, as in *justice department official*. The only occurrences of this type found in our corpus before 1800 were proper names with multiple titles, as in *Minister Count Kinski*, *Lieutenant Colonel Longueville*, *Madame Countess d'Etrees*. NNN sequences begin to appear in common noun phrases in the late nineteenth and early twentieth centuries, but they are still relatively rare; examples in the ARCHER Corpus include: *army reorganization scheme*, *cancer research*

associations/fund, Dublin hospital reports, home rule bill, interest charge amounts, river colony politics, trade union leader.

The dramatic change in the use of these structures occurred in the second half of the twentieth century, when NNN sequences become relatively common, and even NNNN sequences are not unusual. Consider the following examples:

1950–1990 Newspaper prose:

Three-noun sequences:

air force machines, aviation security committee, fighter pilot training, health department clinics, house personnel office, justice department intervention, justice department official, justice department spokesman, news agency correspondent/reports, oil tanker drivers, police motorcycle outriders, polio vaccination situation, road haulage association/drivers/industry, settlements tax increases, task force officers, trade boycott campaign, transport labour unions

Four-noun sequences:

emergency cabinet committee meetings, peace treaties enforcement action

1950–1990 Medical prose:

Three-noun sequences:

acid phosphatase activity/levels/test/units/values, air flow limitations, artery blood flow, assay dilution factor, blood glucose level, blood pressure clinic, body surface area, chromosome gene product, daytime serum concentrations, granulocyte adhesion functions, granulocyte surface membrane, haemoglobin digestion method, hazards regression analyses, hepatitis surface antigen, hill committee report, hill report recommendations, hospital record departments, infarction blood pressure, insulin infusion tests, life insurance tables, (early) morning urine specimens, mouse ascites fluid, nitrogen excretion supply, peak plasma concentrations/levels, Pearson correlation coefficients, plasma concentration curve, plasma glucose levels/profile/value, rabbit immunoglobulin fractions, (high) resolution image intensifier, sinus node dysfunction

Four-noun sequences:

life table survival curves, peak mean plasma concentration, mean plasma glucose value, plasma concentration time curve

Even longer NN sequences are attested in recent decades, including the following sequence of six pre-modifying nouns from the headline of a recent science news article published in the *Guardian* (November 14, 2014):

Philae comet lander alien 'cover-up' conspiracy theories emerge

E) The meaning relationships between the pre-modifying noun and the head noun

Finally, there has been historical extension in the meaning relationships that hold between a pre-modifying noun and the head noun (see Feist 2012 for a detailed discussion of the order of constituents used as NP pre-modifiers). The title nouns and place nouns found as pre-modifiers in the seventeenth and eighteenth centuries are straightforward in this regard. For example, *Captain Smith* is a person named Smith who is a captain; *Greenwich Park* is a park located in Greenwich; *frontier garrisons* are garrisons located at the frontier.

The other classes of pre-modifying nouns, though, do not consistently express a single meaning relationship to the head noun. This is true even for the concrete/tangible nouns found as pre-modifiers in the eighteenth century. For example:

<u>NN sequence</u>	<u>Meaning paraphrase</u>
<i>iron chain/tools, linen handkerchief</i> <i>milk diet, chamomile tea, flannel</i> <i>roller, mineral spring(s)/ poisons</i>	an N2 that is composed of N1
<i>hen eggs, goose eggs, fountain</i> <i>water, spring water</i>	an N2 that comes from N1
<i>coffee house, copper mill, corn field,</i> <i>gun ship, sand bank</i>	an N2 where one can find N1
<i>sea captains, sea commissioners</i>	an N2 specializing in N1
<i>cannon ball, farm house, town wall</i>	an N2 used for/with an N1

The set of possible meaning relationships expands greatly in the late nineteenth century and throughout the twentieth century, associated with the wider range of pre-modifying nouns. Following is a sample of the additional meaning relationships commonly expressed by NN sequences in the late twentieth century:

<u>NN sequence</u>	<u>Meaning paraphrase</u>
<i>government official, union member</i>	a person (N2) belonging to the institution identified by N1
<i>state convention, union assets</i>	an inanimate entity (N2) associated with the institution identified by N1
<i>family history, algebra textbook,</i> <i>sports magazine, psychology lecture</i>	a text (N2) about the topic identified in N1
<i>awards bureau, investigation</i> <i>department, price commission,</i> <i>safety officials</i>	a person or institution (N2) that regulates or administers N1

<i>casualty department, intelligence agencies, news agency, terrorism centre</i>	an institution (N2) that tries to obtain information about N1
<i>extradition treaty, monopoly act, research fund</i>	an inanimate entity (N2) that regulates or administers N1

The grammatical use of nominalizations or nominal conversions as head nouns in these complex noun phrases results in additional meaning relationships. In many cases, the pre-modifying noun is the semantic patient or theme of the process described by the nominalized head noun. Many of these head nouns are derived from intransitive verbs; in those cases, the pre-modifying noun corresponds to the logical subject but the semantic patient. Thus, compare:

Noun phrase: *wage increases*

Clause with intransitive verb: *wages increased*

Clause with transitive verb: *something increased wages*

Other head nouns are clearly derived from transitive verbs; in those cases, the pre-modifying noun corresponds to the logical direct object, expressing the semantic patient. Thus, compare:

Noun phrase: *waste disposal*

Clause with transitive verb: *someone disposed of the waste*

In other cases, the pre-modifying noun identifies the purpose or topical domain of the process described by the nominalized or noun-converted head noun (e.g., *peace conference*). For example:

<i>wage increases, child development, justice department intervention, eye movement</i>	N1 is a patient or theme affected by the process described by N2; syntactically, N1 is the logical subject of N2
<i>weight loss, waste disposal, income tax, taxi driver, trade legislation</i>	N1 is a patient or theme affected by the process described by N2; syntactically, N1 is the logical direct object of N2
<i>trade agreement, reprisal raids, freedom movement, peace conference</i>	N1 identifies the purpose or topical domain of the process described by N2

In some of these cases, both nouns are nominalized processes. For example, a *regression analysis* refers to 'someone analyzing the way in which variable *X* regresses with variable *Y*'.

In addition, there are many more specific meanings for particular NN sequences. For example:

<u>NN sequence</u> <i>age group</i>	<u>Meaning relationship</u> a group consisting of people with a particular age
<i>correlation coefficients</i>	coefficients that report correlations
<i>emergency powers</i>	powers that can be used in an emergency
<i>inoculation experiments</i>	experiments that test the effectiveness of inoculation
<i>retail outlet</i>	an outlet that sells retail merchandise
<i>pressure hose</i>	a hose able to withstand pressure
<i>pressure ratio</i>	a ratio measuring pressure
<u>NNN sequence</u> <i>oil tanker drivers</i>	<u>Meaning relationship</u> people who drive tankers that contain oil
<i>trade boycott campaign</i>	a campaign to encourage people to boycott trade
<u>NNNN sequence</u> <i>emergency cabinet committee meetings</i>	<u>Meaning relationship</u> meetings of a committee associated with a cabinet, called in an emergency
<i>peace treaties enforcement action</i>	activities carried out to enforce treaties intended to result in peace

Summary: Pre-modifying nouns in written discourse These functional extensions in the use of nouns as NP pre-modifiers are further illustrated by considering academic texts from the nineteenth versus late twentieth century. For example, [Text Sample 5.1](#) is from multi-disciplinary science writing (from the *Philosophical Transactions*) in the mid-nineteenth century, illustrating the types of clausal elaboration that were common in earlier centuries (including finite adverbial clauses, complement clauses, and relative clauses). *Of*-genitive constructions and nominalizations (*division, operation, tenacity, explanation, alterations, comparison*) are also common in this text; however, there are no instances of nouns as noun pre-modifiers.

Text Sample 5.1

Nineteenth century multi-disciplinary science

Finite relative clauses are underlined; *of*-genitives are shown in *ITALIC CAPS*.

That division *OF* these nerves produces some serious lesion is proved by the death *OF* the animal, which generally takes place a few days after the operation. Considering the well-known tenacity *OF* life

possessed by these animals this was quite an unexpected result, for which I am unable at present to afford any satisfactory reason... Whatever may be the true explanation, it is impossible not to regard this result with surprise, when we consider the serious lesions which this animal is capable of undergoing at other points OF the frame without loss OF life. The usual time which it survives is variable, and depends greatly on the season OF the year... For the purpose OF avoiding this loss OF life, I adopt the plan OF dividing the glossopharyngeal on one side only OF the tongue, and I find that it has the desired effect OF preserving the life OF the animal, while we can observe the same alterations on the corresponding side, as well as when both nerves are divided. Another advantage found in the division OF a single nerve is, that on the uninjured side we have constantly at hand a means OF comparison by which we can judge with certainty respecting any alterations that may be produced in the divided nerve.

Augustus Waller. 1850.

‘Experiments on the Section of the Glossopharyngeal and Hypoglossal Nerves of the Frog[. . .]’

Philosophical Transactions, 140: 423–429.

Present-day multi-disciplinary science writing can sometimes be similar in grammatical style to nineteenth century science articles. For example, [Text Sample 5.2](#) illustrates a modern *PT* text with extensive clausal elaboration (finite relative clauses are underlined) and frequent use of the *of*-genitive (shown in *ITALIC CAPS*). However, this text also illustrates the recent shift towards increasing use of phrasal modification, with a much greater reliance on nouns as noun pre-modifiers (shown in ***bold italics***).

Text Sample 5.2

Twentieth century multi-disciplinary science

From Darwin derives not only the explicit assumption *OF* ***animal-human continuity***, but also the implicit assumption that ***behaviour patterns***, from which intelligence must be inferred, are proper candidates for evolutionary selection, a tenet that became the ***foundation stone*** *OF* ethology and was also seminal for many branches *OF* psychobiology. In one *OF* his notebooks Darwin wrote that ‘he who understands baboon would do more towards metaphysics than Locke’ (Notebook M, p. 84, 16 August 1838; cf. Gruber (1974) pp. 317–318). Darwin himself, however, is relatively silent, even impotent, when it comes to those behavioural descriptions that might lead to a deep or systematic understanding *OF* the cognitive capacities *OF* animals, let alone metaphysics, and his most concentrated treatment *OF* reasoning, for example, in the Descent *OF* man is little more than scattered anecdotes *OF* doubtful value.

L. Weiskrantz. 1985.

‘Categorization, cleverness and consciousness’

Philosophical Transactions (Series B), 308: 3–19.

Most present-day multi-disciplinary science articles employ much less clausal elaboration, coupled with an extremely dense use of phrasal modifiers, as [Text Sample 5.3](#) illustrates:

Text Sample 5.3

Twentieth century multi-disciplinary science

Tide gauge measurements suggest that global average **sea-levels** rose by between 1 and 2 mm yr^{-1} during the twentieth century (Church et al. 2001). Between 1993 and 2000, **satellite altimetry** indicated that the rate *OF* rise was approximately 2.5 mm yr^{-1} (Cabanes et al. 2001), which may indicate a recent acceleration *OF* the **long-term trend** or it could be the result *OF* interdecadal variability. In the future, global average **sea-level** is expected to increase more rapidly as a result *OF* anthropogenic **climate change**. The Third **Assessment Report** *OF* the Intergovernmental Panel on **Climate Change** concluded that between 1990 and 2100, the global average **sea-level rise** is likely to be between 9 and 88 cm (Church et al. 2001). While this **century-scale rise** in **sea-level** will threaten many low-lying unprotected coastal areas, it is the extremes *OF* **sea-level** associated with **storm surges** that will cause much *OF* the damage (for instance McCarthy et al. 2001). Extreme **water levels** will also increase as **time-average sea-levels** rise. In addition, changes in the number, path and strength *OF* atmospheric cyclonic storms may alter the formation and evolution *OF* **storm surges**. However, because individual **surge events** depend on the driving meteorology, we cannot predict individual surges more than, at most, a few days ahead.

J.A. Lowe and J.M. Gregory. 2005.

'The effects of climate change on storm surges around the
United Kingdom'

Philosophical Transactions (Series A), 363: 1313–1328.

This grammatical style is even more prevalent in recent specialist science research articles, exhibiting an even more extreme reliance on nouns used as noun pre-modifiers, coupled with even less frequent use of dependent clauses. For example:

Text Sample 5.4

Twentieth century specialist science

Population growth rate *is a particularly powerful index for evaluating* **harvest effects** because it measures the ability *OF* a population to increase when subjected to any specified level *OF* exploitation... Selectivity *OF* the harvest on **Putauhinu Island** translates into large differences in **harvest rates** among **weight classes**... The **population effects** *OF* removing an individual depends on quality (i.e. future contributions to reproduction) and on the contribution *OF* its stage to demography. The effect of **quality differences** among **chick classes**

in our model is small because chicks move among classes during the nanao and rama (Table 1). The contribution *OF* **chick survival** to **population growth** is small, and regardless of initial weight all chicks become identical pre-breeders at the end *OF* the first year (Fig. 1)... There is evidence for such links between characteristics *OF* young individuals and **life history traits** *OF* adults in many taxonomies...

Christine M. Hunter and Hal Caswell. 2005.
'Selective harvest of sooty shearwater chicks: Effects of
population dynamics and sustainability'
Journal of Animal Ecology, 74: 589–600.

One interesting characteristic of the pre-modifying nouns found in modern specialist science writing is that they are often *not* technical terms. Of course, we can find counter-examples which fit the stereotype of jargon-heavy scientific academese, as in:

The **Peutz-Jegher syndrome tumor-suppressor gene** encodes a **protein-threonine kinase**.

However, the more common pattern is to modify a non-technical head noun with another non-technical pre-modifying noun. Specialist background knowledge is required to understand the meaning relationship between the two nouns, but the nouns themselves typically have transparent, everyday meanings. Many of these could be paraphrased with an *of*-genitive as nominal post-modifier:

<i>sea level</i>	level <i>OF</i> the sea
<i>water level</i>	level <i>OF</i> the water
<i>life history</i>	history <i>OF</i> a life
<i>life history traits</i>	traits <i>OF</i> a history <i>OF</i> a life

Other examples could be paraphrased with other prepositional phrases as nominal post-modifier:

<i>quality differences</i>	differences in quality
<i>century-scale rise</i>	a rise on the scale of a century
<i>population effects</i>	the effects on the population

In these paraphrased examples, specialist knowledge is required to understand the meaning relationship between the head noun and the prepositional phrase. However, other combinations are even more difficult to paraphrase and thus require even more specialist background knowledge for understanding; for example:

<i>tide gauge measurements</i>	measurements made by a gauge that calculates the level of the tide
<i>long-term trend</i>	a trend that lasts for the long term
<i>time-average sea-levels</i>	the level of the sea averaged across different time periods
<i>storm surge</i>	a surge in the sea level, associated with a storm
<i>surge events</i>	events that involve a surge in the sea level, associated with a storm
<i>harvest rates</i>	differences in the magnitude of a harvest
<i>harvest effects</i>	the extent to which the magnitude of a harvest varies
<i>chick survival</i>	the extent to which chicks survive

As shown from these examples, the individual nouns employed in noun–noun sequences are often not technical, but the meaning relationships underlying those sequences are technical and require specialist background knowledge for understanding. Thus, the historical shift towards an increasing reliance on such forms has resulted in a surprising *loss* of explicitness in specifying the intended meaning. We return to this rhetorical/pragmatic change in [Chapter 6](#).

In sum, the dramatic increase in the use of nouns as nominal pre-modifiers represents much more than a mere quantitative shift in stylistic preference. In particular, the present section has shown that there has been a major expansion in the types of nouns that can occur as nominal pre-modifiers, as well as a major expansion in the range of meaning relationships and grammatical functions underlying noun–noun sequences.

5.2.2 *Attributive adjectives as nominal pre-modifiers*

Attributive adjectives differ from most other types of phrasal noun modifiers in their historical development. As we showed in [Chapter 4](#) (see especially [Table 4.4](#) and [Figure 4.5](#)), attributive adjectives were already well established and frequent in use in the seventeenth century (in popular registers as well as informational registers), and their rate of occurrence has changed little over the subsequent centuries. However, this stability in quantitative patterns masks historical linguistic change in the use of this linguistic feature.

As described in Biber et al. (1999: 508ff), attributive adjectives can be grouped into two major semantic categories: ‘descriptors’ and ‘classifiers’ (cf. Feist 2012: 9–10). Descriptors are prototypical adjectives describing attributes of the head noun, including physical characteristics, evaluations, and emotional states. They are used to describe a wide range of physical

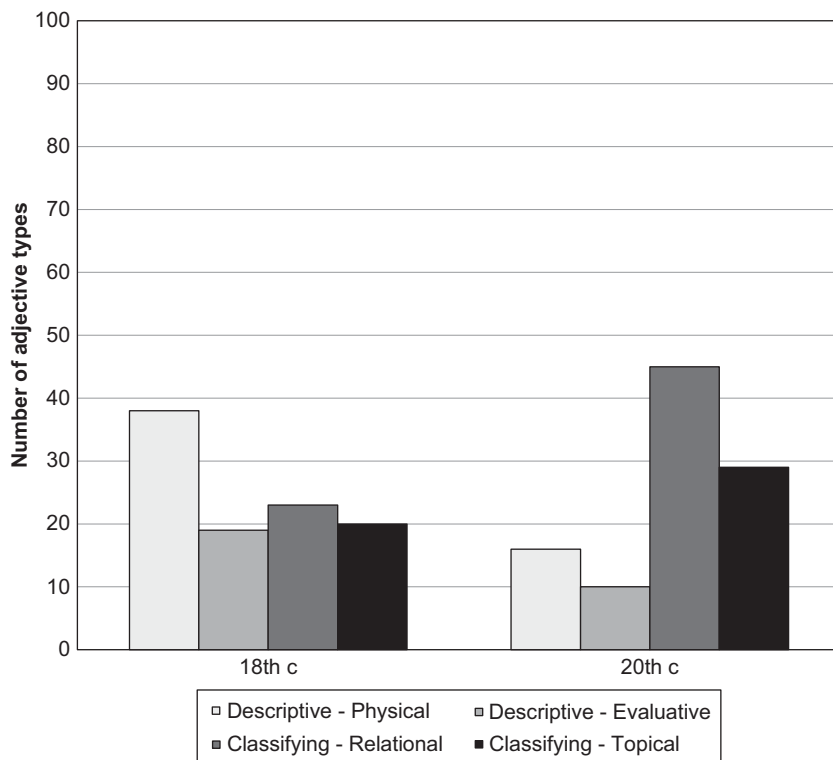


Figure 5.3 The 100 most frequent adjective types, categorized for descriptive versus classifier functions

characteristics, such as color or physical appearance (e.g., *red, pale*), size or weight (e.g., *large, thin, heavy*), age and frequency (e.g., *old, new, daily, recent*). Descriptors are also used to denote a wide range of evaluations or emotional states (e.g., *bad, nice, beautiful, appropriate, strange, happy, sad*).

In contrast, adjectives functioning as classifiers delimit or restrict the reference of the head noun, identifying the category of the noun (e.g., *final, previous, German, chemical, legal*). Classifiers are usually not gradable, while descriptors usually are (compare *very large* versus **very chemical*). There are two major subcategories of classifiers: relational/classificational adjectives, which serve to delimit the referent of a noun in relation to other referents (e.g., *additional, average, different, general*), and topical/affiliative adjectives, which identify the type or subject area of a noun, or designate the group that a referent belongs to (e.g., *commercial, human, official, social, Chinese, European*).

There are major differences across present-day registers in the reliance on descriptor versus classifier adjectives: conversation (and fiction) tend to use descriptors to a much greater extent than classifiers, while newspaper prose and

academic prose tend to use classifiers to a much greater extent than descriptors (see Tables 7.2 and 7.3 in Biber et al. 1999: 511–513). However, in the seventeenth and eighteenth centuries, academic prose was similar to popular registers (e.g., fiction and letters) in that it mostly utilized descriptor adjectives.

To further investigate the extension of attributive adjective functions in academic prose, we coded the functions of attributive adjectives in science research articles from two historical periods: seventeenth/eighteenth century versus twentieth century. One difference that quickly became apparent was the range of different adjectives ('adjective types') used in these two historical periods. That is, although the overall rate of occurrence for attributive adjectives is similar for eighteenth versus twentieth century research writing (see Figure 4.5 in Chapter 4), the number of different adjective types used in these periods is dramatically different, with nearly four times more adjective types in the most recent period. That is, eighteenth century science writing tended to rely on the frequent use of relatively few different adjectives, whereas modern science writing employs a much larger inventory of different adjectives.

Figure 5.3 shows that the typical functions of attributive adjectives in science writing have similarly shifted over time. This figure is based on a categorization of the one hundred most frequent adjectives in eighteenth century versus twentieth century science articles, distinguishing descriptive versus classifying functions. In the eighteenth century, descriptive functions (physical and evaluative) were the most prevalent for this set of adjectives (c. 65% of the top 100 adjectives), while that preference was reversed in the twentieth century (with c. 75% of the top 100 adjectives serving classifier functions). In both historical periods, descriptive adjectives that identify physical attributes (e.g., *great*, *large*, *high*) are almost twice as common as descriptive evaluative adjectives (e.g., *good*, *important*, *special*), although physical descriptors are often used for evaluative functions in academic writing (*a large difference*, *a small effect*, *high ability*). Within the classifier category, relational adjectives (e.g., *same*, *different*, *general*, *major*) are more common than topical/affiliative adjectives (e.g., *social*, *political*, *public*), and this relationship also holds for both historical periods.

When we consider adjective types that occur with lower frequencies in academic science writing, we see that topical/affiliative adjectives dominate these lists, especially in the twentieth century. For example, in a sample of less common attributive adjectives (occurring c. one time per million words), 50% served topical functions (e.g., *perimental*, *cultural*, *ellipsoidal*, *coniferous*), compared with c. 25% with relational functions (e.g., *middle*, *eligible*, *existent*, *contrary*). Surprisingly, descriptive-evaluative functions are much more important among these lower frequency adjectives than descriptive-physical functions: 20% of these adjectives serve descriptive-evaluative functions (e.g., *famous*, *urgent*, *erroneous*), compared with only 5% with descriptive-physical functions.

In sum, while there has been little historical change in the overall frequencies for attributive adjectives, there have been important changes in the diversity of adjective types, and in the typical functions of those forms. In particular, present-day science writing employs a much wider set of different adjectives than in previous centuries, with most of those new adjectives serving classifying-topical functions.

5.2.3 Noun-participle compounds as nominal pre-modifiers

Noun-participle compounds (e.g., *correlation-based category*) are another phrasal feature associated with present-day academic writing. As noted in Chapter 4, noun-participle compounds were rarely used in eighteenth century academic writing. While they are still not especially frequent in absolute terms, they have increased in use by a factor of ten during the last half century (see Figure 4.6). This increase has been especially noteworthy for N + *-ed* compounds (e.g., *cue-focused processes*, *mood-induced states*, *stress-induced anisotropy*).

Similar to the grammatical features described in the preceding sections, noun-participle compounds were comparatively restricted in function in earlier historical periods (in addition to being infrequent). By the mid-nineteenth century, several relatively lexicalized compounded forms were already in use, such as *home-made*, *self-made*, *star-spangled*, *frost-bitten*, *heart-broken*, *lion-hearted*, *fun-loving*, *blood-curdling*, and *awe-inspiring*. Certain words were repeatedly incorporated into these compounds. For example, numerous compounds were formed with the prefixed nouns *self* and *heart* (e.g., *self-made men*, *self-imposed discipline*, *self-taught artist*, *self-organized vigilance*, *self-destroying evils*, *self-denying Jesuit*, *self-relying men*, *self-sacrificing traits*; *heart-broken Incas*, *heart-rending cry*, *heart-cheering salute*, *heart-chilling scenes*, *heart-sickening scene*). Similarly, the participles *shaped* (e.g., *heart-shaped muscle*, *spindle-shaped globules*, *spoon-shaped extremities*) and *colored* (e.g., *saffron-colored powder*, *coffee-colored discharge*) were used to form multiple compounds in texts from this period.

Beyond those relatively fixed forms, nineteenth century authors used noun-participle compounds to express a range of relatively concrete meanings, combining nouns that refer to physical entities with participles expressing tangible actions. These included both compounds formed with *-ed* participles (e.g., *the surf-beaten shore*, *blood-drenched plains*, *storm-beaten trees*, *ivy-covered towers*, *the ice-choked Delaware River*) as well as compounds formed with *-ing* participles (e.g., *flesh-eating barbarians*; *money-squandering community*; *nostril-dilating power*; *fur-bearing animals*; *a health-seeking, pleasure-loving man*; *ship-building operations*; *salt-producing carbonate*; *air-breathing reptile*).

Many of these noun-participle compounds that were used in eighteenth and nineteenth century academic writing have now become relatively

informal, being used regularly in modern popular spoken and written texts (e.g., fictional novels and even conversation regularly use forms like *home-made*, *self-made*, *heart-broken*, *fun-loving*, *awe-inspiring*). However, innovative uses of this grammatical construction in present-day English are restricted primarily to informational academic prose, and in that register, the lexical associations of this feature have been extended greatly from the nineteenth century.

In modern science and social science research articles, the two most common participial forms incorporated into noun-participle compounds are *based* and *related*. Interestingly, these are both abstract and relational in meaning, in contrast to the tangible actions/events described by the compounded participial forms in the nineteenth century. In addition, most of the nouns incorporated into these structures are abstract in meaning, in contrast to the nouns referring to physical entities in the seventeenth and eighteenth centuries. Examples include:

space-based observations

matrix-based analysis

XML-based modelling

GT2-based services

grid-based mechanisms

performance-based measures

object-based competition

standards-based tests

age-related increases

health-related problems

education-related income disparities

science-related skills

accountability-related pressures

time-related covariates

instruction-related activities

the PTdins(3,5)P2-related fission cascade

three Rb-related genes

systemic sclerosis-related symptoms

Beyond the participial forms *based* and *related*, there is an extremely diverse set of nouns and participles incorporated into noun–particle compounds in present-day science and social science research writing. As observed earlier, most of these involve abstract nouns combined with process, stative, or relational participles. For example:

service-oriented architecture

disease-associated Prpres

GPI-linked proteins

protein-lined canaliculus

stress-induced anisotropy

HPC-enabled simulations

climate-driven changes

atom-resolved evidence

scrapie-infected mice

fluorescence-activated cell

river-supplied sediment

basin-distributed temperature

HIV-infected individuals

placebo-controlled design

investigator-designed questionnaires

teacher-reported behavior problems

life-threatening conditions

number-crunching simulations

light-scattering measurements

voltage-sensing apparatus

ligand-binding domains

protein-signaling networks

population-averaging effects

sediment-gauging activities

mountain-draining rivers

colony-forming units

answer-changing behavior

Thus, noun–participle compounds have undergone an extension in functionality similar to that of the other devices used for the pre-modification of nouns in academic written English. Most obviously, that extension has involved a large increase in the inventory of words used to form these compounds. However, this extension also entails a shift in the kinds of nouns and participles that can be combined to form a compound. In the eighteenth and nineteenth centuries, these words were restricted primarily to concrete physical objects (expressed by the nouns) and concrete actions/events (expressed by the participles). In contrast, the noun–participle compounds found in present-day academic writing usually consist of abstract nouns combined with an abstract process/relational participle.

5.3 Phrasal features functioning as post-modifiers of a head noun

5.3.1 *Prepositional phrases as nominal post-modifiers*

In [Chapter 4](#), we described the strong decline in frequency for finite relative clauses and *of*-genitive phrases in academic prose over the last century. At the same time, though, ‘other’ prepositional phrases (e.g., phrases headed by *in*, *on*, *for*) functioning as noun modifiers strongly increased in frequency. Specifically, [Figure 4.8](#) showed that ‘other’ prepositional phrases functioning as noun modifiers nearly doubled in frequency in academic prose from 1850 to 2005. Post-nominal prepositional phrases headed by two specific prepositions have been especially important in this historical change: *in* and *for* (see [Figure 4.9](#)). Post-nominal phrases headed by *in* are especially prevalent in science research writing, while post-nominal phrases headed by *for* are more common in the social sciences. For example:

*The presence of ingested particles **IN** the multivesicular bodies **IN** the relatively early stage confirms the findings of Farquhar et al. . .* [Science writing]

*Whenever reference to the city as eternal was made it was associated with a concept **FOR** the need **FOR** Roman social and military unity.* [Social science writing]

However, this grammatical change extends well beyond the increased use of those two prepositions. In fact, the most notable finding here is the large number of different prepositions that function as the head of prepositional phrases modifying nouns, reflecting the extent to which this structural device has been

generalized in modern academic writing. Six other prepositions (in addition to *in* and *for*) occur with moderately high frequencies heading post-nominal phrases in modern academic prose: *about*, *between*, *from*, *on*, *to*, and *with*:

Coders were unaware in regard to other information **ABOUT** the families.

We examine the relationships **BETWEEN** class and country. . .

Items **FROM** the Home Inventory were used to construct two composite measures. . .

The literature **ON** cultural congruence has focused primarily on the interactional and discursive aspects of teaching.

Immediately after exposure **TO** wasps, pupae were weighed. . .

Nop7p has been implicated in the control of ribosome biogenesis via its interaction **WITH** origin recognition complex proteins. . .

In addition, a wide range of other prepositions are used at lower frequency levels for this function in modern academic writing, including *across*, *after*, *against*, *along*, *among(st)*, *around*, *as*, *at*, *before*, *behind*, *by*, *into*, *like*, *near*, *over*, *toward(s)*, *under*, *upon*, *within*, and *without*:

This cooperation **ACROSS** the aisle was sparked by FDR's controversial "court-packing" scheme.

In late 1939, a few weeks **AFTER** the outbreak of the war, a group of scientists persuaded the financier Alexander Sachs to convey to Roosevelt a letter. . .

Settlement expansion also took place . . . at other sites **AROUND** Jerusalem.

Puerto Rico's status **AS** a commonwealth of the United States creates a unique set of Conditions. . .

We used corrected-item-mean imputation to handle missing data **AT** the item level. . .

Closer teacher-child relationships predicted higher standardized test scores in language and math and higher ratings **BY** teachers. . .

Until shortly before his entry **INTO** politics, Barak had been a military man. . .

Monuments occasioned a struggle **OVER** meaning **AMONG** social groups. . .

[moderation could be achieved] by taming the radical forces **WITHIN** the Nazi Party. . .

Thus, this structural feature (prepositional phrases as nominal post-modifiers) is currently in advanced stages of being generalized to the full

set of prepositions, rather than being restricted to only a small subset of specialized prepositions.

Academic research writing often incorporates sequences of these prepositional phrases as noun modifiers, creating dense information structures with few verbs. In the following example, four of the noun phrases in the sentence contain prepositional phrases as noun post-modifiers, including multiple modifiers for a single head noun, and prepositional phrase modifiers embedded in other modifiers. For example:

*Here we will use translocation **OF** muskox *Ovibos moschatus* **FROM** eastern Greenland **TO** the mountain range Dovre **IN** Central Norway to parameterize a simple population model **WITHOUT** density regulation that can be used to improve our ability to predict future fluctuations **OF** small populations.*

Thus, at one level, modern academic prose is similar to eighteenth century science writing in employing extensive structural embedding. The main difference, though, is that the embedding in modern texts is phrasal, while the embedding in eighteenth century texts is clausal.

Similar to the historical changes for nominal pre-modifiers, the increased frequency of prepositional phrases as nominal post-modifiers has been accompanied by a large expansion in function and meaning. The most obvious indication of this expansion is the increase in the range of meaning relations that can be signaled by prepositional phrase (PP) noun modifiers. For example, in present-day academic prose, the prepositions *in* and *on* can express both concrete locative meanings (*blood flow in skin and in skeletal muscle; many places on the slopes below*) and abstract meanings (*evolution in Roosevelt's imperialist thought; the influence of evolutionary ideas on his public policies*). However, the use of these prepositions to express abstract meanings is a relatively recent development restricted primarily to writing. Thus, in conversation, over 90% of these PPs as noun modifiers express concrete/locative meanings, as in:

somebody in the neighborhood

a vault in his house

those kids in California

stuff on the dash

that thing on the roof

did you read that dedication on the first page

This is similar to the typical uses of these prepositions in written texts from the sixteenth to the seventeenth centuries, where they normally expressed concrete/locative meanings even in academic written prose. Consider the following examples.

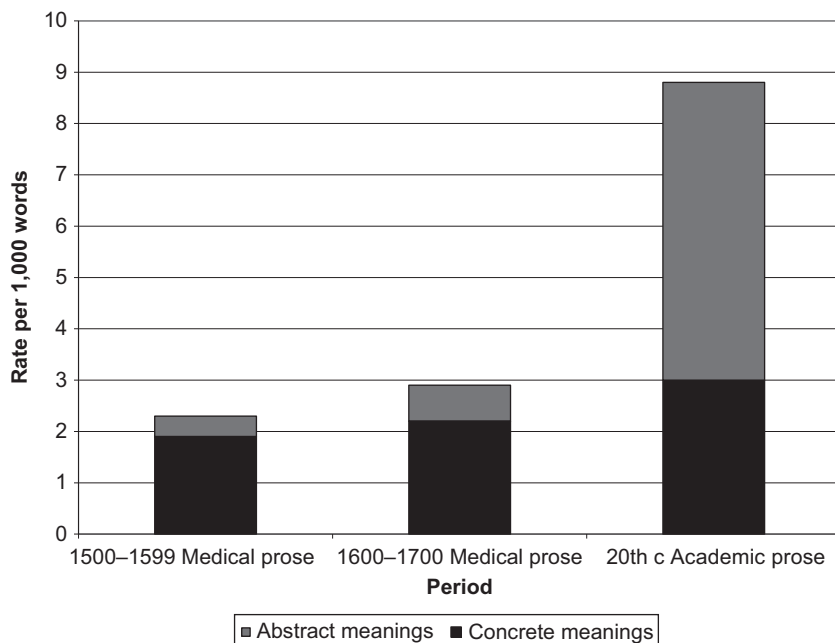


Figure 5.4 IN as noun modifier: concrete versus abstract meanings

Concrete meanings expressed by noun-modifying PPs with *in*; seven-teenth century medical prose:

<u>N + PP sequence</u>	<u>Meaning</u>
<i>a wynde in the heed, ache in the backye, pain in his knee</i>	location inside a body part
<i>the oil in the thermometer, kernels in your meat, quantity of opium in it</i>	location inside an object or substance
<i>our apothecaries in England, John Bissite in St. Peters Parish</i>	geographic location
<i>the foregoing chapters in the first part, his judgement and candor in his writings</i>	textual location

Concrete meanings expressed by noun-modifying PPs with *on*; seven-teenth century medical prose:

<u>N + PP sequence</u>	<u>Meaning</u>
<i>a postume on the longes, blisters and hackes on the lips, two on each side</i>	location on the surface of an object

As Figures 5.4 and 5.5 show, these concrete meanings were dominant for *in/on* noun-modifying PPs in the sixteenth and seventeenth centuries.

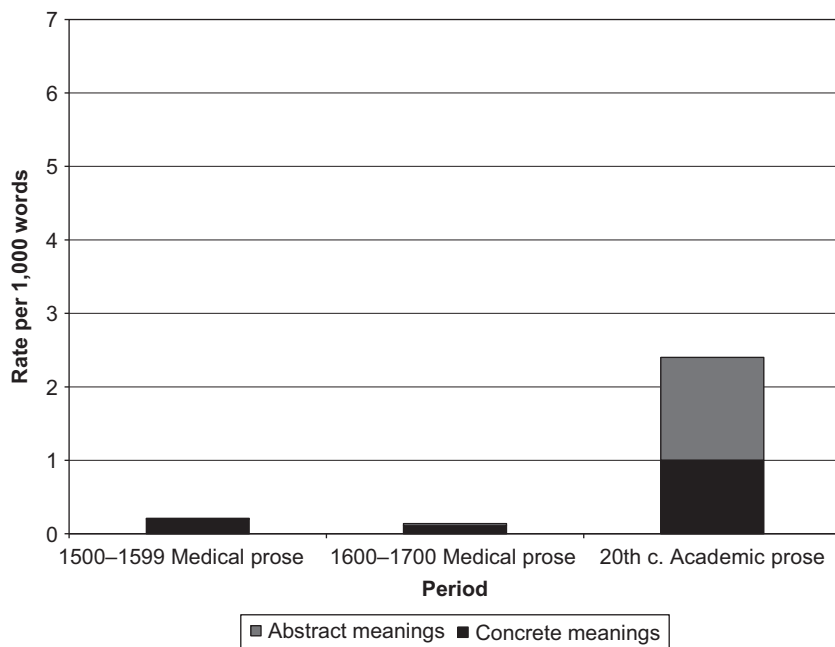


Figure 5.5 ON as noun modifier: concrete versus abstract meanings

In contrast, abstract meanings were uncommon. Abstract meanings with *on* were especially rare, restricted to an identification of ‘topic’, as in:

a poem on the virtue of a laurel leaf

some remarks on the late debate about X

occasional notes on Dr. George Thompsons

a learned author on this subject

PPs with *in* had a similar abstract use in the seventeenth century, identifying a general research domain:

my studie in ciuile and humane learnynge

that Axiome in philosophie

his learning in all sciences

our work in the Astrological or chymical way of physyc

new discoveries in medicine

my sentiments in these matters

Figures 5.4 and 5.5 show that the functions of both *in* PPs and *on* PPs as noun modifiers in academic writing has changed dramatically from the seventeenth century to the present day: in present-day academic prose, c. 60% of these constructions are used to express abstract meanings rather than concrete/locative meanings. Those abstract uses include the identification of topical domain, as in seventeenth century medical prose. However, there are other abstract uses that were only beginning to emerge in the seventeenth and eighteenth centuries. Three of these abstract functions are especially important in modern academic science writing:

1. nominalization + *on* PP, corresponding to a nominalized equivalent of a prepositional verb construction; for example:

abiological dependence on physical conditions

[compare: it depends on physical conditions]

his reliance on the evolutionary thesis

[compare: he relied on the evolutionary thesis]

a focus on measures of student outcomes

[compare: someone focused on measures of student outcomes]

2. process noun + *on* PP, corresponding to a transitive verb with a direct object that is the 'patient' of the verb; for example:

a profound influence on student responses

[compare: something profoundly influenced student responses]

A significant impact on sea-bird numbers

[compare: something impacted sea-bird numbers]

a much greater effect on the attitudes of other students

[compare: something affected students' attitudes]

3. constructions with an *ing* non-finite clause following the preposition *on*; for example:

an effect on determining choice

emphasis on providing support

any restriction on publishing it

research on promoting science learning

Prepositional phrases with *in* have also developed two especially common abstract uses in modern academic writing:

1. process noun + *in* PP, corresponding to an intransitive verb with a subject that is the 'patient' of the verb; for example:

a decline in mean IQ

[compare: mean IQ declined]

a sudden increase in metabolic rate

[compare: metabolic rate suddenly increased]

a change in the food supply

[compare: the food supply changed]

2. constructions with an *ing* non-finite clause following the preposition *in*; for example:

invaluable assistance in recording electrocardiographs from canaries

difficulty in separating the sarcoplasmic proteins from the myofibrils errors in rounding the total score

One of the more interesting aspects of these historical changes with *in* and *on* PPs is that each of these prepositions has developed a specialized abstract function that marks the modifying noun as the semantic 'patient' of the process described by the head noun. For example:

an increase in efficiency

[compare: something increased efficiency, or efficiency increased]

influence on dropout rates

[compare: something influenced dropout rates]

For PPs with *in*, these head nouns usually correspond to intransitive verbs:

Subject NP ('patient') + Intransitive V

> Head noun (identifying a process) + *IN* + NP ('patient')

the chief reason for Britain's decline in exports

[compare: Britain's exports declined]

a rapid increase in the size of the egg

[compare: the size of the egg increased]

variation in frequency

[compare: the frequency varied]

For PPs with *on*, these head nouns correspond to transitive verbs. Many of these are simple transitive verbs, as in:

Transitive V + Object NP ('patient')

> Head noun (identifying a process) + *ON* + NP ('patient')

a significant influence on the tragic developments that followed

[compare: something influenced the tragic developments]

two factors have the greatest impact on college grades

[compare: two factors impact college grades]

a greater emphasis on intellectual behaviors

[compare: someone emphasises intellectual behaviors]

Other head nouns taking *on*-PPs correspond to prepositional verbs, although the object of the verb still serves the same semantic role of ‘patient’:

Transitive prepositional V + *on* + Object NP (‘patient’)

> Head noun (identifying a process) + *ON* + NP (‘patient’)

a biological dependence on physical conditions

[compare: it depends on physical conditions]

his reliance on the evolutionary thesis

[compare: he relied on the evolutionary thesis]

a focus on measures of student outcomes

[compare: someone focused on measures of student outcomes]

In present-day informational writing, there are numerous head nouns that can take these patterns. Seven head nouns are especially frequent with prepositional phrases with *in* followed by an NP (‘patient’): *change*, *decrease*, *difference*, *fall*, *increase*, *rise*, and *variation*. (Other head nouns that occur with this pattern include: *advances*, *breakdown*, *decline*, *development*, *evolution*, *growth*, *improvement*, *reduction*, and *shift*.) Similarly, there are numerous head nouns that take *on* plus an NP (‘patient’): *effect* and *emphasis* are especially frequent with this pattern, but other nouns include *attack*, *constraint*, *debate*, *decision*, *discussion*, *impact*, *influence*, *limit*, *limitation*, *restriction*. There are fewer prepositional verbs that have been nominalized to occur as head nouns in this pattern, but examples include *bearing on*, *dependence on*, *focus on*, *insistence on*, and *reliance on*.

These specialized grammatical uses have been introduced gradually over the past four centuries in informational written texts. The semantic ‘patient’ pattern with the preposition *in* historically precedes the ‘patient’ pattern with *on*, corresponding to the more general development of *in* and *on* PP as nominal modifiers (see Figure 4.9). *In* PPs with this meaning seem to have first occurred in the sixteenth century with the head nouns *difference* and *change*, and this pattern then gradually spread to other process nouns during the following four centuries. We list here the first attested examples for selected process nouns + *in* + NP (‘patient’) from the OED:

- 1530 no **difference in** sounde
- 1585 to attempt alteration and **change in** the church of God
- 1627 one inch of **decrease in** the growth of men
- 1760 improvement in **the art of ... defending strong places**
- 1804 a **fall in** the price of the article
- 1806 great **variation in** its composition
- 1817 a small **rise in** the annual payment
- 1823 a sudden **increase in** the circulating medium
- 1869 quick **growth in** intellectual and moral force

As noted earlier, the development of *on*-PPs as abstract noun modifiers lagged behind *in*-PPs. The phrase *dependence on* (derived from the prepositional verb *depend on*) occurs at a relatively early stage, but nominalized equivalents of other prepositional verbs with *on* appear considerably later. We list here the first attested examples of these combinations from the OED:

- 1605 **dependence on** the Latin
- 1754 **reliance on** the promises of God
- 1798 **insistance on** tradition
- 1908 a **focus on** passing pedestrians

Simple process nouns that take *on* + patient NP (i.e., nouns that are not derived from prepositional verbs) lagged behind these other structures by about a century; for example:

- 1696 Absence, Madam, has had the same **effect on** my Passion
- 1869 their **impact on** the atoms
- 1899 their **emphasis on** the partial nature of all physiological analysis.

Another major reflection of functional expansion is the increasing ability of all prepositions to occur as a nominal post-modifier with an *-ing* clause as complement. In essence, these structures function to express a process in a nominal structure, while at the same time retaining clausal constituents in the *ing*-clause, such as objects and adverbials. For example:

- the first step **in seeking quicker treatment**
- difficulty **in separating the sarcoplasmic proteins from the myofibrils**
- errors **in rounding the total score**

the value of full doses in treating cancer

a book on managing large organizations

a critical study on reducing hardware development time

emphasis on providing support

motives for committing a violent crime

plans for creating a UFO

procedures for selecting cluster structure

Early occurrences of this structural pattern are attested from Middle English and Early Modern English; the following examples are from seventeenth century medical prose:

the vertues and worth of this Medicine in helping and curing many diseases

pains in making water

the difficulty in searching out the causes of them

the truth of the marchaunt in transporting the same

As shown in Figure 5.6, this structure was rarely used during the seventeenth to nineteenth centuries, but it has increased in use over the course of the twentieth century. Prepositional phrases with *in* and *for* take the lead here, while phrases with *on* + *-ing* clause are expanding in use more slowly.

It is possible to track the historical progression of this structure to an increasing number of different controlling nouns. In many cases, these structures have an earlier historical counterpart with a prepositional verb + *-ing* clause. The prepositional verb later becomes nominalized (or converted to a noun), but the structure retains the co-occurrence with an *ing*-clause as complement. Thus, compare the following examples from the *OED*, showing the first attestation of a verb + *in* + *-ing* clause versus the corresponding noun + *in* + *-ing* clause:

VERB: 1572 *such as men use in searching ore*

versus

NOUN: 1620 *it hath special use in illustrating [something]*

VERB: 1647 *assist in procuring [something]*

versus

NOUN: 1742 *assistance in apprehending [something]*

VERB: 1730 *succeed in forcing [something]*

versus

NOUN: 1820 *success in capturing [something]*

VERB: 1606 ***persist*** *in doing wrong*
versus
NOUN: 1874 ***persistence*** *in seeking a pacific settlement*

VERB: 1940 [they] ***specialize*** *in hunting big game*
versus
NOUN: 1970 ***specialisations*** *in printing [something]*

Interestingly, only abstract meanings occur with post-modifying *in* followed by an *-ing* clause. This is in part due to the non-finite verb phrase coming after the preposition, as it represents a process rather than a concrete entity. However, this is also due to the abstract nature of all head nouns that occur with this pattern (e.g., *accuracy*, *difference*, *effect*, *degree*, *value*, *confidence*, *error*, and *diligence*). As the following examples illustrate, many of these nouns are nominalizations, and the meaning relationship between the head noun and the process represented in the prepositional phrase + *-ing* clause is one of simultaneous time or place. That is, the abstract state (usually

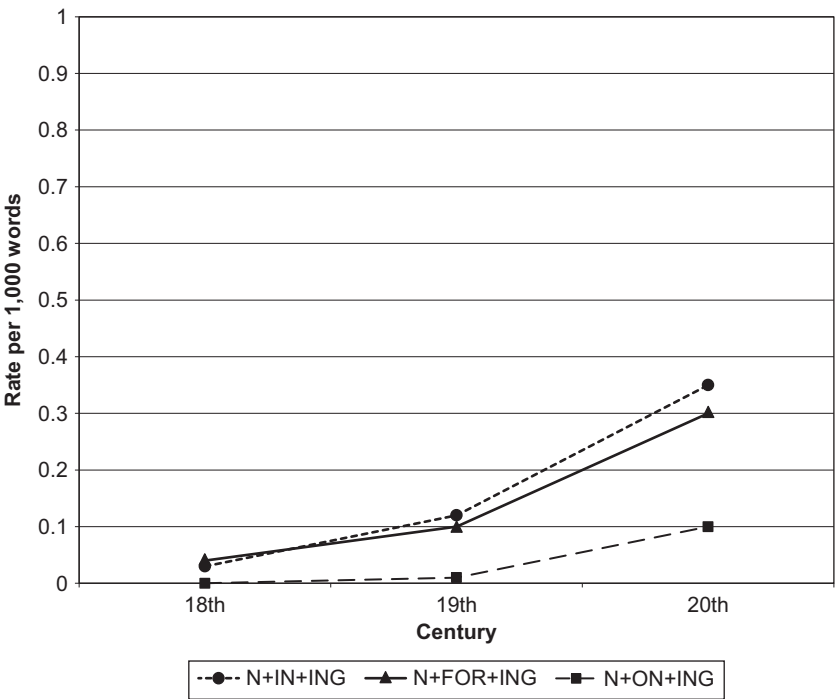


Figure 5.6 Noun + preposition + *ing*-clause in academic prose

expressing a stance or evaluation) referred to by the head noun exists at the same time and/or place as the process in the verb phrase.² For example:

Boldness in foretelling the Emperors death (1702)

[compare: boldness when he foretold the Emperors death]

the greater ease and convenience in solving all the necessary geographical and nautical problems (1777)

[compare: ease and convenience when he solves all the...problems]

his diligence and exactness in making observations for a long series of years (1786)

[compare: diligence and exactness when he made observations]

difficulty in conceiving how this is possible (1811)

[compare: difficulty when he (tries to) conceive how this is possible]

difficulty in expanding his religious conceptions (1858)

[compare: difficulty when he (tries to) expand his religious conceptions]

In post-modifying *-ing* structures with the preposition *for*, two major types of meaning are expressed. In the first, *for* + *-ing* clause modifies head nouns like *equation*, *machine*, or *formula* – nouns that are somewhat more concrete and tangible in meaning than cognitive abstract nouns like *idea*, *reason*, or *caution*. These *for*-phrases function to identify the purpose of the entity referred to by the head noun, and they can usually be paraphrased with a relative clause like *'which function(s) to'* or *'which serve the purpose of'*. For example:

the Rules for directing Ships into Ports through the vast ocean (1715)

[compare: rules that direct ships]

equations for reducing the mean Place of the 4th Satellite (1749)

[compare: equations which function to reduce]

² A second pattern with the preposition *in* was relatively common in the eighteenth and nineteenth centuries, but appears to have become obsolete by the late twentieth century. In this pattern, the prepositional phrase follows a noun of time (e.g., *year*, *minute*, *hour*), and the relationship between the noun and the prepositional phrase seems to be one of conveying the duration of the process referred to by the *-ing* verb. Thus, in the first example given here, 'he is 29 years in moving through his orbit' could be paraphrased as 'he takes 29 years to move through his orbit'.

Time reference + in + V-ing

he is 29 **years** in moving through his orbit (1767)

the Particles of Light are 8 **Minutes** in passing the Distance from the Sun to us (1779)

she differs but two **hours** in rising for six days together (1811)

and differs only 28, 24, 20, 18 or 17, **minutes** in setting (1811)

and being about thirty **years** in performing this revolution (1817)

a **machine** *for subdividing the arc of this quadrant* (1790)

[compare: a machine for the purpose of subdiving]

a **formula** *for finding the dominical letter in any century* (1822)

[compare: a formula which functions to find the dominical letter]

the **data** *for computing the distance of the heavenly bodies* (1860)

[compare: data for the purpose of computing the distance]

operations *for determining the figure of the earth* (1893)

[compare: operations for the purpose of determining the figure of the earth]

A second less common meaning expressed by *N + for + -ing* clause is more abstract: the head noun refers to the evidentiary basis of an argument, and the *for*-phrase presents the argument (or conclusion, hypothesis, etc.) itself:

good **reasons** *for thinking that the adjustment may be performed* (1790)

[compare: reasons to think that the adjustment may be performed]

good **grounds** *for concluding that such was the case* (1889)

[compare: grounds to conclude that such was the case]

The functional extension of noun-modifying PPs headed by ‘other’ prepositions could be described in much more detail. We have focused here primarily on the development of *in/on/for* as noun-modifying PPs, but similar extensions in function have occurred for many of the other prepositions that head post-nominal PPs. The descriptions in this section have documented linguistic innovation and change at multiple levels: in the overall extent to which a structure is used; in the particular structural variants that occur; and in the meaning relationships expressed by the structure.

5.3.2 *Appositive noun phrases as nominal post-modifiers*

Chapters 3 and 4 show that appositive noun phrases are common in informational writing but extremely rare in conversation. Examples like the following are commonplace in present-day newspaper prose (head noun phrase **bolded**, appositive noun phrase underlined):

“A lot of companies are looking to get out of these plans anyway,” says

Dan McGinn, a pension consultant in Anaheim, Calif.

The only comparable feature that occurs in conversational discourse is a sequence of a noun with vague reference followed by a more precise noun phrase, as in:

*You can turn on **the thing**, the Tony awards.*

These ‘prefaces’ and ‘tags’ in conversation are used by speakers to clarify the meaning of noun phrases. They are attached to the beginning or end of a clause, rather than modifying a noun phrase within a clause (see Biber et al. 1999: 1072ff); for example:

That little shop – it’s great.

*He’s had **a blind** put up – a special blind that leads straight across the fanlight.*

In contrast, appositive noun phrases modify a head noun, and the entire complex noun phrase structure is embedded in a clause. For example:

*Last month the nine pages at issue were leaked, apparently by someone within the military, since **the Pentagon memo**, an internal document, was attached.*

As Figure 4.10 (in Chapter 4) showed, it is only in recent decades that appositive noun phrases have become common, even in informational written discourse. These structures were attested but rare in the eighteenth century, but they have increased in use considerably over the past one hundred years.

A more detailed consideration of these structures shows that they have also expanded greatly in function, similar to the patterns of use seen for NN sequences and PPs as noun modifiers. In the eighteenth century, most appositive noun phrases were used to identify the title or position of a person, as in the following examples.

Newspaper:

1773: On Friday last arrived here from London, via New-York, **Mr. William Young**, Botanist to their Britannick Majesties

1785: The church of Seaham having some time ago been broke into, and robbed of the communion plate, **Mrs. Robinson**, relict of the late Rector of that place, has generously repaired the loss by a handsome set, of the value of 50l.

1819: The King on Thursday granted an audience of leave to **M. de NESSELRODE**, the Russian Minister for Foreign Affairs who had been for some time at Paris, though his name has not till now been once mentioned in the Journals since his arrival there.

Science prose:

1775: The bounty of **his Majesty**, our Patron, happily removed this difficulty.

1825: At the period of my transmitting my former paper to the Society, I was not aware that any English author had written upon the same subject; but was afterwards much surprised on learning that **a Gentleman of Edinburgh**, Mr. DALYELL, had published an account of these animals in 1814,...

This function continues to be one of the most common uses of appositive noun phrases in modern newspaper writing, as in:

James Klein, president of the American Benefits Council

Dallas Salisbury, CEO of the Employee Benefit Research Institute

However, additional functions began to develop in academic research writing in the nineteenth century. First, appositive noun phrases began to be used to provide additional descriptive information about the head noun. In many examples, the first NP is a common noun referring to a person, and the second NP gives descriptive characteristics of that person. The following examples are all from mid-nineteenth century science/medical writing:

their daughter, a young servant girl in the seventh month of pregnancy

the mother of the girl, a strong healthy woman, aet. 45, and nearly at the full period of pregnancy

a previously healthy girl of five years, a patient of Dr Menzies

the patient, a young girl about 13 years of age

This pattern of use was further extended in the nineteenth century to provide descriptive information about inanimate head nouns, as in:

The first case of erysipelas, the one from which all the other seizures start, occurred at Corfu. . .

It is the American or green helebore, and is found from Canada to Georgia. **The root**, the only efficient part, is perennial, bulbous, stout, premorse, and dark.

A specialized function introduced around this time is to explain a technical term with an appositive noun phrase, providing a common-place description of the term:

or still more strictly in relation to **Favus**, a disease dependent on the existence of a plant or vegetable like parasite-phyticides

a parasitic affection frequently styled ringworm of the body,- to distinguish it from **Tinca (Herpes) Tonsurans**, ringworm of the scalp

In addition, a structural/textual innovation was introduced in the nineteenth century: the use of parentheses to mark the appositive noun phrase, rather than separating the two by commas. For example:

In about two months after it had acquired this additional head, a fragment separated from **the tail** (the most usual place of separation) and was in progress towards its entire reproduction when it was accidentally lost. . .

the former is composed of sporules, **empty tubes** (the mycelium), and tubes filled with sporules

Despite this variation in form and meaning, appositive noun phrases throughout the nineteenth century retained the primary meaning relationship of being directly co-referential with the head noun. Thus, in the examples mentioned earlier, *Mr. William Young is Botanist to their Britannick Majesties*; *the patient was a young girl about 13 years of age*; *the root is the only efficient part*; *tinca (herpes) tonsurans is ringworm of the scalp*, and so on.

This directly co-referential meaning relationship continues to be the primary one expressed by appositive noun phrases in present-day newspaper prose. However, in twentieth century academic research writing, appositive noun phrases began to be used to express information having a wider range of implicit meaning relationships to the head noun. These noun phrases can be quite long and complex, and so they are usually contained in parentheses, rather than separated by commas.

Often, an appositive noun phrase is used to introduce an acronym, or to itemize the members of a group, as in:

*We present the results of the **International Meta-analysis of Mortality Impact of Systemic Sclerosis (IMMISS)**. . .*

*In **four cohorts** (Athens, Keio, Mayo, and Florence), investigators stated that. . .*

However, in many other cases, the meaning relationship is more specialized in academic research writing. For example, consider the following sentence, where the appositive noun phrase is not strictly co-referential with the head noun:

*Comparison of these scores to the studies in our meta-analysis reveals that they are all of **moderate quality** (scores of 2 to 4 on a scale of 0 to 5).*

In this sentence, the head noun gives the author's interpretation of the findings ('moderate quality'), while the appositive noun phrase provides the actual results.

Many other instances of appositive noun phrases in academic writing illustrate even more distant relationships. Thus consider:

*Numerous variables were measured, including case status, sex, race, **date of enrollment** (date of first visit to the cohort with the pertinent diagnosis), age at first visit, . . .*

In this example, the head noun identifies a variable that was used in a study, and the appositive noun phrase is used to specify methodological information, documenting how that variable ('date of enrollment') was measured in the study. The meaning could have been made more explicit through the use of clausal structures, something like:

**For the purposes of this study, we operationally defined the variable 'date of enrollment' as the date when the subject made the first visit to the cohort with the pertinent diagnosis.*

Appositive noun phrases are commonly used for such functions in modern scientific writing. The following example illustrates the dense use of appositive noun phrases in a science research article, often with appositives embedded in other appositives:

*Timing of peak growth coincides with **the observed peak growth of our wild parr**; a period when **their growth effectiveness** (assimilation minus BMR) was high (fig. 1d) and the period when their natural food was particularly abundant.*

The highest level appositive in this example expresses the classical co-referential meaning: *the observed peak growth of our wild parr is a period when their growth effectiveness was high*. However, the embedded appositive noun phrase (*assimilation minus BMR*) has a more specialized meaning that provides details about the methodology, specifying that 'growth effectiveness' was calculated by subtracting the BMR measure from the assimilation measure.

Appositive noun phrases with complex meaning relationships are common in modern science research articles. In all cases of apposition, there are no grammatical markers used to make the meaning relationship explicit. Thus, consider the following examples:

*Analyses that included **all cases in each center** (n = 3311; total follow-up: 19,990 person-years) yielded largely similar results.*

***Our Girnock analysis** (fig. 2a) shows that **late autumn and winter** (day 240 onwards) was the only period when...*

*In multivariate analyses that adjusted for age and sex, **renal involvement** (hazard ratio [HR] = 1.9; 95% confidence interval [CI]: 1.4 to 2.5), **cardiac involvement** (HR = 2.8; 95% CI: 2.1 to 3.8), **pulmonary involvement** (HR = 1.6; 95% CI: 1.3 to 2.2), and **anti-topoisomerase I antibodies** (HR = 1.3; 95% CI: 1.0 to 1.6) increased mortality risk.*

As these examples show, it is not at all unusual to have multiple appositive noun phrases in the same clause, often modifying a single head noun, and often embedded hierarchically. For example, 'HR' and 'CI' are embedded appositive noun phrases providing acronyms for 'hazard ratio' and 'confidence interval'. These appositive noun phrases all provide additional descriptive information about the head noun, but many of them do not express a simple co-referential relationship. From a discourse perspective, these devices

provide a way for authors to condense information into few words. We turn to this more general functional priority in the following section.

5.4 The systemic 'drift' towards structural compression, motivated by economy of expression

The grammatical developments documented in the last two chapters represent a fundamental shift in the discourse style of academic writing, resulting in grammatical patterns of use that are unlike those found in any register in earlier historical periods. In the present chapter, we have shown how these changes entail more than just stylistic shifts in the density of nominal/phrasal features; rather, these grammatical features have also undergone important extensions in their uses and functions.

The underlying functional motivation shared by these historical changes is the drive towards economy of expression, resulting in structural compression to convey the maximum amount of information in the fewest words possible. From a cultural perspective, these linguistic developments have occurred alongside the proliferation of academic sub-disciplines, which have become increasingly specialized in both topic and readership, resulting in the 'information explosion' and the need to present more information in an efficient and concise way.

When considered from this general functional perspective, the grammatical features discussed in preceding sections can be treated as a cline rather than as a simple dichotomy between clausal versus phrasal devices. At the most general level, all phrasal devices are more structurally compressed than clausal devices. However, there is also variation within these two major categories. For example, within the general category of clausal modifiers, non-finite clauses are more compressed than finite clauses. Similarly, within the general category of phrasal modifiers, NP pre-modifying phrases are more compressed than NP post-modifying phrases.

Thus, the structural devices used for noun phrase modification can be ranked along a cline of compression as follows:

Least compressed

- 
- Noun phrases modified by finite relative clauses
 - Noun phrases modified by non-finite relative clauses
 - Noun phrases modified by post-modifier phrases (PPs, appositive NPs)
 - Noun phrases modified by pre-modifier phrases (attributive adjectives, pre-modifying nouns)

Most compressed

Up to this point in our discussion, we have for the most part treated these as independent structures, describing the functions of each grammatical device separately from the others. However, in many cases, these devices are interchangeable, and thus they can be regarded as structural alternatives for expressing essentially the same information. Thus, consider the following sentence from an eighteenth century science article, focusing on the underlined noun phrases:

And if his Computation, which was made for Greenwich, had been reduced to the Meridian of London, the Difference would have been still less.

This same information could easily be expressed with a non-finite relative clause rather than a full finite relative clause:

And if his Computation made for Greenwich had been reduced to the Meridian of London, the Difference would have been still less.

However, this structure can be further compressed to a simple prepositional phrase as post-modifier, as in:

And if his Computation for Greenwich had been reduced to the Meridian of London, the Difference would have been still less.

And finally, the two noun phrases with PP post-modifiers could be further reduced to structures with a noun as pre-modifier, as in:

And if his Greenwich Computation had been reduced to the London Meridian, the Difference would have been still less.

This cline of compression provides a framework for considering the entire system of grammatical complexity, and historical trends within that system. When all of these grammatical features are considered together, they show that the system of grammatical complexity has been undergoing a ‘drift’ towards greater structural compression, with phrasal pre-modifiers of nouns representing the most extreme endpoint of that trend.

Biber and Finegan (1989a) document the historical trend towards ‘oral’ discourse styles in popular written registers (letters, fiction, essays), associated with the increasing use of a suite of colloquial grammatical features. That study employed the concept of historical ‘drift’ to describe ‘a cumulative series of gradual linguistic developments in a consistent direction’ (Biber and Finegan 1989a: 489). Sapir (1921: 155) introduced this concept as follows: ‘The drift of a language is constituted by the unconscious selection on the part of its speakers of those individual variations that are cumulative in some special direction’. In the 1989 Biber and Finegan study, and in other studies of ‘informalization’ (Fairclough 1992) or ‘colloquialization’

(Hundt and Mair 1999; Mair 2006; Leech et al. 2009), that 'drift' relates to the use of a wide range of conversational grammatical features, which have steadily increased in frequency in popular written registers during the twentieth century.

The grammatical changes described in the present book are equally systematic, and they similarly involve an entire system of grammatical features moving in 'a special direction'. However, this historical 'drift' is exactly the opposite from the one described in previous research: towards increased structural compression associated with a distinctive 'literate' discourse style (rather than change towards increased use of colloquial features).

Thus, Figure 4.7 traced the dramatic decline in the frequency of finite relative clauses in academic prose over the last 300 years, and Figure 4.8 traced the similar decline in the use of prepositional phrases as adverbials and the decline in the use of *of*-genitives. In contrast, Figure 4.5 showed the dramatic increase in the use of nouns as NP pre-modifiers, especially in the last century. Likewise, Figure 4.6 showed the recent increase in the use of noun-participle compounds as NP pre-modifiers; Figures 4.8 and 4.9 plot the increase in use for 'other' prepositional phrases as NP post-modifiers; and Figure 4.10 showed the strong increase in the use of

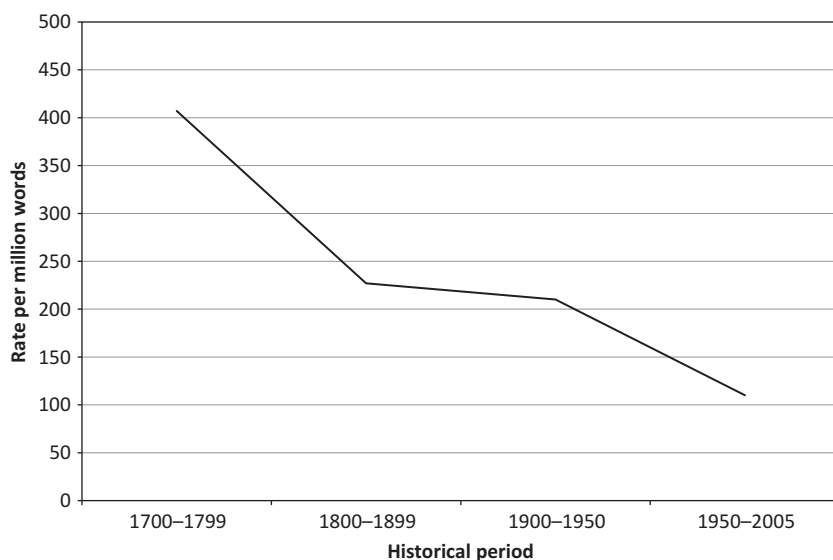


Figure 5.7 Historical change in the use of predicative relative clauses in science research articles
(i.e., finite relative clauses with a subject gap and the main verb BE)

appositive noun phrases as NP post-modifiers. Taken together, these plots document the overall systemic change in the use of grammatical complexity features in academic written registers: the drift towards a structurally compressed discourse style.

Similar to the example sentence discussed earlier, it is possible to treat these changes for individual grammatical features as structural alternatives, exploring how similar kinds of content can be expressed through increasingly compressed structural devices. As a case study, consider the use of finite relative clauses that have a simple predicative function, with a subject gap and the main verb *be*, as in:

I took away the glass with the water, which was quite limpid

Figure 5.7 plots the historical use of such relative clauses, showing a historical trend similar to that for finite relative clauses overall (compare Figure 4.7 in Chapter 4) – except that the decrease in frequency for predicative relative clauses is even greater than for relative clauses generally.

Consideration of specific examples from the eighteenth century shows that most predicative relative clauses could be easily rephrased in modern prose with phrasal noun modifiers. Consider the following examples:

Predicative relative clauses that could be replaced with attributive adjectives:

A body that is light will swim
[compare: a light body]

I took away the glass with the water, which was quite limpid
[compare: the quite limpid water]

The most powerful diuretics and hydragogue purgatives were given by another physician and me in vain, the bulk of her belly, which was very great when I saw her first, rather increasing
[compare: the bulk of her great belly]

I cut a little of the upper end of the branch which was very tender
[compare: the very tender branch]

Predicative relative clauses that could be replaced with a post-modifying prepositional phrase:

It was their custom to name these climates from some considerable place that was at or near the middle.
[compare: some considerable place at or near the middle]

Its motion represents the apparent diurnal motion of the heavens, which is from east to west

[compare: the apparent diurnal motion of the heavens, from east to west]

The water which is in E B, is by the Forcer D C forced up, whereby the Air, that is between A and E, is compressed. . .

[compare: the water in E B]

[compare: the Air between A and E]

A Gentleman in the Country, sitting by his Fire-side in the Winter, was swelled for several Days with the Smoak or Flame of some Poyson-Wood that was in the Fire.

[compare: some Poyson-Wood in the Fire]

Predicative relative clauses that could be replaced with a post-modifying appositive noun phrase:

This species of air (which is a kind of acid) is a modification of the nitrous acid. . .

[compare: This species of air (a kind of acid)]

He supposeth therefore that Sirius, which is the brightest of all the Fixed Stars, may be equal to the Sun

[compare: Sirius, the brightest of all the Fixed Stars]

I passed into the little hole a sprig of a known Plant, which was Baulme. . .

[compare: a known Plant (Baulme)]

If the short line VA be imagined perpendicular to the plane of the ecliptic, and to pass through the planet at [A], the angle VTA, is the latitude of the planet, which is called the geocentric latitude

[compare: the latitude of the planet (the 'geocentric latitude')]

As noted above, phrasal modifiers can be ranked among themselves for the extent to which they are structurally compressed, with NP pre-modifiers being more compressed than post-modifiers. Interestingly, there is some evidence that the overall historical drift towards more compressed structures in academic prose extends to even this level, with phrasal post-modifiers recently being replaced by phrasal pre-modifiers.

For example, in [Section 5.2.1](#), we discussed how pre-modifying nouns can be considered as a structural alternative to post-modifying *of*-phrases. It turns out that many of the head nouns that permit this choice can also occur with 'other' prepositional phrases as post-nominal modifiers

(see Section 5.3.1). By focusing specifically on the use of those head nouns, it is possible to track more general historical change in the preferred structure used in cases when a range of variants is possible.

For example, several head nouns commonly occur with *on* prepositional phrases, including: *effect*, *emphasis*, *attack*, *constraint*, *debate*, *decision*, *discussion*, *impact*, *influence*, *limit*, *limitation*, and *restriction* (see Section 5.3.1). Most of these same head nouns can occur with an *of*-genitive, although the specific meaning usually differs: the noun following *of* identifies the semantic agent of the activity described by the head noun, while the noun following *on* identifies the semantic patient of the activity described by the head noun. The two prepositions often occur together in the same sentence with these meanings, as in:

*His surprize will give place to concern and regret, when he observes [...] the pernicious **influence of** uncleanliness and foul air **on** the duration of life*
[compare: uncleanliness influences the duration of life]

*[This. . .] visibly explains the **effect of** a severe colic **on** the intestine*
[compare: a severe colic affects the intestine]

Pre-modifying nouns can also modify these same head nouns, often with an agentive meaning; they thus often serve as direct structural alternatives to *of*-genitives with these head nouns. For example:

*Observational learning may produce the declarative knowledge **influence on** social decision making.*
[compare: the influence of declarative knowledge on social decision making; i.e., declarative knowledge influences social decision making]

*Because of the CES **emphasis on** local interpretations [...], the larger research project focused on two questions . . .*
[compare: the emphasis of the CES on local interpretations; i.e., the CES emphasizes local interpretations]

Figure 5.8 plots the historical development of all three grammatical devices controlled by this set of head nouns (based on analysis of the Corpus of Historical American English – COHA; see Davies 2012). The two features that can be regarded as direct structural alternatives – *of*-genitives and pre-modifying nouns – are almost mirror images of each other, with pre-modifying nouns increasing in use over the decades at about the same rate as the decrease in use of *of*-genitives. The third structural option – *on* post-modifying phrases – follows an

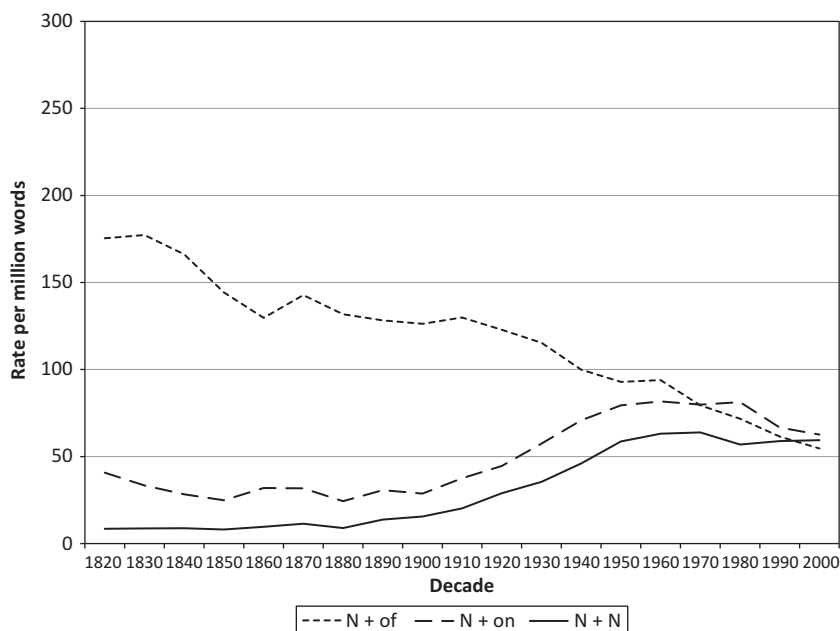


Figure 5.8 Historical change (in COHA) of three grammatical variants:

N* + of, N* + on, N + N*

Head nouns [N*]: *effect, emphasis, attack, constraint, debate, decision, discussion, impact, influence, limit, limitation, restriction*

increasing trajectory that parallels pre-modifying nouns. That grammatical feature represents the emergence of a new grammatical/pragmatic function: the noun phrase following *on* identifies the semantic patient of the process referred to by the head noun. Thus, we see two quite different patterns of historical change here: *of*-genitives are being replaced by pre-modifying nouns (a more compressed structure) that serve a similar function, while post-modifier *on*-phrases are independently increasing in use because they serve a function not attested in earlier periods.

A similar kind of analysis can be carried out for the head nouns that commonly control *in*-phrases, including: *change, decrease, difference, fall, increase, rise, variation, advances, breakdown, decline, development, evolution, growth, improvement, reduction, and shift*. These head nouns also regularly occur with both *of*-genitives and pre-modifying nouns. However, in this case all three structural variants usually have equivalent meanings.

Thus, consider the following sentences from eighteenth century science articles, compared with the rephrased expressions that would be more typical in present-day prose:

A change of velocity will not alter the elongations

[compare:

A change in velocity will not alter the elongations

A velocity change will not alter the elongations]

The duration is calculated in proportion, allowing for the difference of refraction in reference to the different state of the atmosphere at different seasons.

[compare:

difference in refraction

refraction difference]

Figure 5.9 plots the historical change in the use of these three grammatical devices controlled by a set of head nouns expressing ‘change’ meanings (based on analysis of COHA; see Davies 2012). Similar to Figure 5.8, we see a steady decline in the frequency of *of*-genitives controlled by these nouns. *In*-phrases were already established with this function in 1820, and they steadily increased in use during the following decades. By the mid-twentieth century, *in*-phrases have overtaken *of*-genitives controlled by these head nouns. However, we then see a surprising development in the latter half of the twentieth century: *in*-phrases decline in use, actually paralleling the declining trajectory of *of*-phrases.

We believe that this reversal in the trajectory of *in*-phrases can be explained by the historical development of pre-modifying nouns with these head nouns. Figure 5.9 shows that noun + noun sequences occurred with low frequency throughout the nineteenth century, but then steadily increased in use over the decades of the twentieth century. In the 1980s, pre-modifying nouns with these ‘change’ head nouns overtook *of*-genitives in frequency, and then in the 1990s, this structure overtook *in*-phrases in frequency. Thus, it appears that the maximally compressed structural alternative – pre-modifying nouns – has become the preferred variant by the end of the twentieth century.

To further confirm this development, we undertook the same analyses in the Corpus of Contemporary American English (COCA; see Davies 2010), allowing a more detailed consideration of recent change during the years between 1990 and 2012. Figure 5.10 confirms the cross-over between

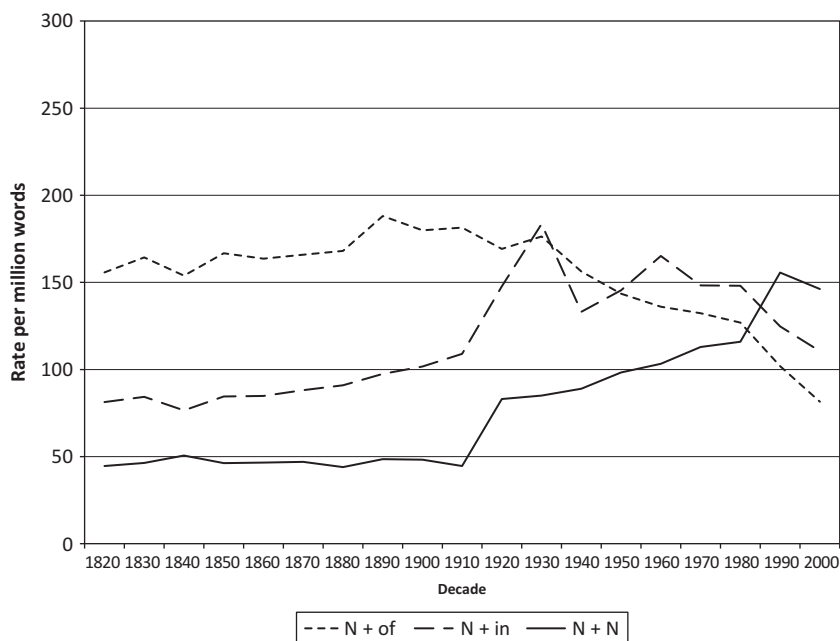


Figure 5.9 Historical change (in COHA) of three grammatical variants:
N* + of, N* + in, N + N*

Head nouns [N*]: *change, decrease, difference, fall, increase, rise, variation, decline, growth*, etc.

pre-modifying nouns and *in*-phrases in the 1990s, and further shows that this trend is still in progress up to the present day.³

In fact, the most dramatic increase in frequency for pre-modifying nouns has occurred in the last five years. As a result, pre-modifying structures like the following are common in present-day academic prose:

Nouns that commonly pre-modify 'change' head nouns in 2005 science and social science articles:

NOUN + <i>growth</i> :	<i>ability, achievement, adolescent, body, cell, individual, juvenile, plant, population</i>
NOUN + <i>change</i> :	<i>attitude, behavior, individual, mood, weight</i>

³ COCA and COHA differ in terms of corpus composition (with a different representation of registers), and thus COCA has a higher proportion of academic texts than COHA. This corpus composition difference is reflected in the apparent jump in frequency for all three variants observed when directly comparing Figures 5.9 and 5.10.

NOUN + <i>difference</i> :	<i>ability, age, gender, mean, standard deviation, (raw) score, test</i>
NOUN + <i>variation</i> :	<i>individual, group, size, weight</i>
NOUN + <i>reduction</i> :	<i>weight, (class) size</i>
NOUN + <i>development</i> :	<i>adolescent, adult, attitude, cancer, career, child, community, competence, concept, curriculum, language, literacy, materials, memory, personality, program, self, staff, teacher, test, vocabulary</i>

Such evidence provides strong confirmation for the general pattern of historical development in academic prose that we have described in the preceding sections: a steady drift in preferred use that affects the entire set of compressed, phrasal structures. These historical trends are relatively continuous from a quantitative perspective, showing a steady decline in use for more elaborated structures, and a steady increase in use for more compressed structures. But in addition, we have shown in the present

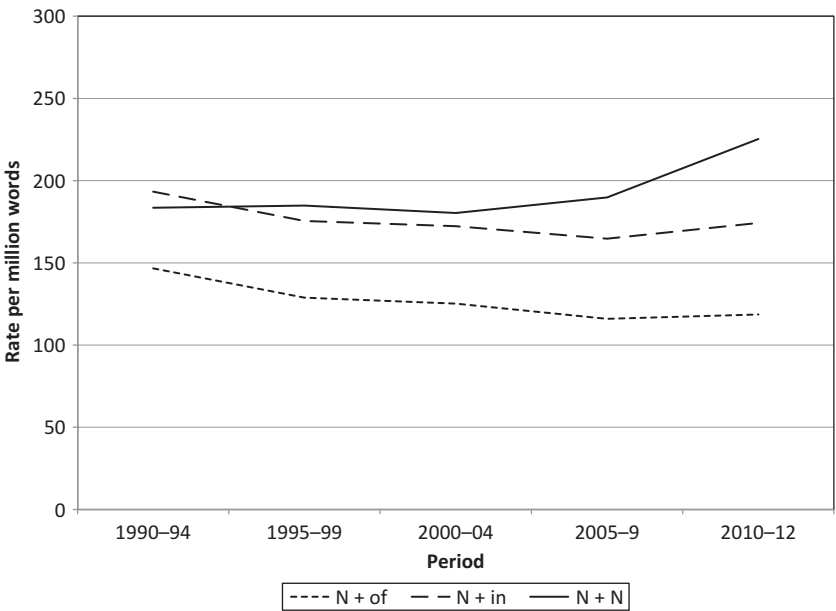


Figure 5.10 Historical change (based on COCA) of three grammatical variants from 1990–2012: *N* + of*, *N* + in*, *N + N**
Head nouns [*N**]: *change, decrease, difference, fall, increase, rise, variation, decline, growth*, etc.

section how these trends operate in concert along a linguistic cline, with the most compressed structures becoming increasingly preferred over time.

We have noted repeatedly how this pattern of inter-related structural changes has resulted in a discourse style unattested in earlier historical periods. In contrast to common stereotypes about academic writing, this dominant discourse style of science research writing is not verbose and needlessly elaborated. In fact, it is maximally compressed in structure.

It further turns out that the grammatical style of science research writing directly contradicts preconceived notions about the communication of information in academic prose. That is, instead of being maximally explicit in the ways in which it conveys meaning, the compressed grammatical features typical of science prose result in a discourse style that is highly inexplicit in meaning. We turn to those considerations in the following chapter.

The loss of explicitness in academic research writing

6.1 Stereotypes concerning explicitness in academic research writing

Researchers have often claimed that academic writing is explicit in meaning, especially in contrast to the vague and implicit expression of meaning in conversation. That is, while speech is dependent on a shared situational context, academic writing is claimed to be ‘decontextualized’, ‘autonomous’, and ‘explicit’ (see, e.g., DeVito 1966; Olson 1977; Kay 1977; Johns 1997: 58–64). This ideal is also commonly advocated by writing instructors, who urge students to overtly encode all assumptions and logical relations in the written text. There is little possibility of direct interaction between the writer and a reader of an academic written text, and thus essentially no possibility for clarification if the meaning is not clear and explicit from the text itself. As a result, academic writing is purported to be maximally explicit in the expression of meaning.

In Chapter 1 (see Section 1.2.2), we documented the dominant and pervasive view that academic writing can be characterized as employing an elaborated and explicit discourse style. These characterizations rely on terms such as *complexity*, *elaboration*, *specificity*, and *clarity* both to describe published academic writing and to identify desired characteristics in student academic writing. It would be nearly impossible to read current writing research and not come away with a general belief that academic writing is ‘complex’, ‘elaborated’, and ‘explicit’.

Functional linguists have made related claims regarding the use of more or less explicit grammatical variants. In particular, Rohdenburg (1996: 151) argues that linguistic variation is constrained by the ‘complexity principle’: ‘in the case of more or less explicit grammatical variants, the more explicit ones tend to be preferred in cognitively more complex environments’. The more explicit variants are more elaborated structurally, with additional grammatical function words or morphemes. For example, the *to*-infinitive

is regarded as more explicit than the bare infinitive; complement clauses with the complementizer *that* are more explicit than complement clauses with a 'zero' complementizer; and finite complement clauses are regarded as more explicit than non-finite complement clauses.

Rohdenburg approaches cognitive complexity from the perspective of syntactic environment. For example, occurrence in a passive voice construction would be regarded as more cognitively complex than occurrence in an active voice construction. The central claim underlying the 'complexity principle' is that increased cognitive complexity of the syntactic environment encourages the use of more explicit grammatical variants.

However, by extension it might be expected that the increased cognitive complexity associated with the informational focus of written academic texts would also result in the use of more explicit grammatical variants. And this is exactly the relationship claimed in many previous descriptions of academic writing (see discussion in [Section 1.2.2](#)).

In the preceding chapters, we have presented strong evidence to challenge the stereotype that academic research writing is structurally elaborated. In fact, just the opposite is the case: academic prose – especially science research writing – relies heavily on extremely compressed grammatical structures, and it has evolved steadily over the last century to favor compressed structures over more elaborated alternatives.

It turns out that there has been a concomitant historical shift towards increasingly *inexplicit* expression of meaning, further contradicting preconceived stereotypes. That is, in direct contradiction to the proscriptions of writing instructors and the expectations of writing researchers, the preferred grammatical structures of academic research writing are much less explicit in meaning than alternative, more elaborated structures. There can be no doubt that written academic texts are more cognitively demanding and complex than most conversational or popular written texts. However, despite this increased cognitive load, written academic texts systematically prefer compressed structural variants, resulting in a dense use of grammatical structures that are *inexplicit* in meaning on a scale that has not been generally appreciated. We explore these meaning relationships in detail in the following sections.

6.2 Explicitness of reference in conversation versus academic writing

It is uncontroversial that spoken conversation is often *inexplicit*, with conversational participants frequently employing expressions with vague and/or implicit references. Conversational participants share the same

physical place and time, they often share extensive personal knowledge about each other, and they have the opportunity to interact and thus clarify any misunderstandings. As a result, spoken discourse relies heavily on pronouns, adverbs, definite noun phrases, ellipsis, and deictic expressions that refer directly to the situational context (or to the participants themselves). The meaning of these forms is not explicit, because the words have no referential meaning outside of a particular situational context. Consider the following example (pronouns are underlined; **situated place adverbials in bold**):

- KATE: See the pictures of the house?
 PETE: No. Where are they?
 KATE: They're over **there on the chair**.
 PETE: Oh, ok.
 KATE: That envelope **right there** honey **where the camera is**. **Under that envelope**.
 PETE: Where did you go? The camera shop?

Even full noun phrases are often situation-dependent in conversation. Thus, based only on the written transcript, we are not able to identify the specific reference of '*the house*,' '*that envelope*,' '*the camera*,' and '*the camera shop*'. However, these noun phrases all have specific references that were easily understood by these conversational participants, who shared the same physical situation and background information.

Situated reference of this type is virtually non-existent in professional academic writing. The reader almost never shares time and place with the author of a written academic text. However, other types of situated reference are present in academic writing. Deictic pronouns can be used to refer to the academic text itself, or to the propositions put forward in the text. Thus, consider:

The doubling time for the DNA of coliphage T2 is about 2 minutes, compared to 20 minutes or more for that of the host cells; for polio-virus RNA it is 20 to 30 minutes, for polyoma about 1 hour, compared to 18 to 24 hours for their host cells. **This** shows that the initiation for nucleic acid synthesis occurs independently for the cellular and the viral nucleic acid.

Renato Dulbecco. 1965.
 'Characteristics of virus-cell complexes'
American Journal of Medicine, 38: 669–677.

In this example, the text itself provides a context, and so pronouns can be used to refer directly to a part of that text.

However, imagine the difficulties for reader comprehension if an author referred directly to the actual time and place where he/she was writing. So,

for example, imagine that one of us pointed to a graph in a book that was sitting on the desk, and we wrote ***This** shows that pronouns are more common in conversation than in writing*. Unless we had previously referred to ‘the graph’ in the text itself, readers would have no clue what *this* referred to. This characteristic holds generally for referring expressions in written texts: if the author fails to use explicit referring expressions, readers will have no chance of identifying the intended reference of those expressions. From that perspective, academic writing is much more explicit in meaning than conversation.

However, an additional consideration is the way in which meaning relations are expressed among grammatical constituents, and in this respect, academic writing is anything but explicit. In fact, it has changed historically to strongly prefer grammatical styles that are dramatically less explicit than in earlier periods.

For example, nominalizations and passives entail reduced explicitness because they omit certain structural elements (see Halliday 1979; Halliday and Martin 1993). Thus, compare:

- a) *The Department of Water and Power manages hazardous waste >>*
- b) *hazardous waste is managed >>*
- c) *hazardous waste management*

In the passive construction (b), we no longer know who the ‘agent’ is. In the nominalized construction (c), it is not even explicitly stated that an activity is occurring. Nominalizations are also inexplicit about the time reference of activities. The clausal constructions in (a) and (b) are marked for tense (compare *someone manages hazardous waste* and *hazardous waste is managed*), and they have the potential to be marked for aspect (compare *someone has managed hazardous waste* and *hazardous waste had been managed*). However, there is no possibility of expressing tense and aspect in the nominalized version (c); the time reference is implicit and must be inferred by the reader.

The loss of explicit meaning associated with nominalizations and passives has been discussed in previous studies by Halliday (see, e.g., Halliday 1979; Halliday and Martin 1993). However, what has been less noticed is that there is a large set of compressed structural devices that behave in similar ways: they are all prevalent in academic writing, and they all result in less explicit meanings. Nominalizations and passives are part of this overall pattern, but there are numerous other grammatical devices that are even more common and similarly inexplicit. In particular, the many forms of phrasal (as opposed to clausal) modification described in the preceding chapters all

result in a loss of explicitness. In the following sections, we describe the lack of explicitness associated with the most important of those features.

6.3 Inexplicit meaning relationships associated with phrasal pre-modifiers in noun phrases

From a quantitative perspective, many of the most dramatic historical changes documented in the preceding chapters involve the increased use of phrasal NP pre-modifiers in academic writing (see, e.g., [Figure 4.5](#)). Nouns as pre-modifiers of a head noun are the most common of these features. Such structures are ubiquitous in present-day informational prose, and as a result, it is easy to gloss over the inexplicit meaning relationships among the elements of these constructions. However, it turns out that these phrasal pre-modifiers represent an extremely wide range of meaning relationships to the head noun, with no overt grammatical marker as to which relationship applies to a particular noun–noun combination. As a result, we could almost characterize these structures as being maximally *inexplicit* in meaning: exactly the opposite of the stereotypes concerning what we would expect in academic writing.

For example, consider the meaning relationship between a head noun and a pre-modifying noun, as in the noun phrase *heart disease*. There are no grammatical clues to help the reader know what the meaning relation is between the two nouns. The surprising fact is that there is an extremely wide range of meanings possible between the two nouns, and thus there is no automatic way to determine the relationship between the two nouns. Rather, the reader must rely on background knowledge (often expert knowledge associated with specific specializations) to infer the relationship between the two nouns. Thus consider:

<i>heart disease</i>	a disease located in the heart
<i>prison officials</i>	officials who work in a prison
<i>union assets</i>	assets that belong to a union
<i>age group</i>	a group of people who are the same age
<i>casualty department</i>	the department that deals with casualties
<i>emergency powers</i>	powers that are in force during times of emergency
<i>income tax</i>	a tax based on your income
<i>monopoly act</i>	an act to prevent monopolies

<i>peace conference</i>	a conference that attempts to establish peace
<i>press conference</i>	an interview (or conference) of a celebrity or spokesperson with newspaper reporters
<i>price commission</i>	the commission that regulates prices
<i>sector strike</i>	a strike involving a particular sector of the economy
<i>sex differences</i>	differences between the two sexes
<i>television interview</i>	an interview that is broadcast on television
<i>time interval</i>	the interval, measured in time, between two events
<i>computation time</i>	the time required to compute something

As a result of this inexplicitness, the exact same pre-modifier can have a completely different meaning relation to the head noun, as in:

<i>pressure hose</i>	a hose able to withstand pressure
<i>pressure ratio</i>	a ratio that measures pressure

Many of these noun phrases have head nouns that refer to a process or event. These nouns are converted from a corresponding verb (e.g., *increase*, *loss*, *offer*) or they are nominalized forms of a verb (e.g., *calculation*, *treatment*, *resemblance*). In many cases, the pre-modifying noun corresponds to the semantic patient of the process or event referred to by the head noun. For example:

<i>wage increases</i>	something increases wages
<i>weight loss</i>	something or someone loses weight
<i>the population decrease</i>	the population decreases; that is, something causes the population to become smaller
<i>teacher training</i>	someone trains teachers
<i>population averaging</i>	someone averages the population statistics
<i>concentration adjustment</i>	someone adjusts the concentration
<i>fault displacement</i>	something displaces the fault

<i>alcohol consumption</i>	someone consumes alcohol
<i>the 3He absorption</i>	something absorbs the 3He
<i>parameter calculations</i>	someone calculates the parameters
<i>drug resistance</i>	someone/thing resists drugs

However, these sequences with a converted/nominalized head noun are also commonly used to express other meaning relations. In a few cases, the pre-modifying noun identifies the semantic agent of the action in the head noun:

<i>a government offer</i>	the government offers something
<i>a Treasury decision</i>	the Department of Treasury decides something

In several other cases, the pre-modifying noun identifies some kind of adverbial in relation to the event/process of the head noun, often with a temporal or locative meaning. For example:

<i>the spring decrease</i>	the decrease that occurs in spring
<i>preshift training</i>	training that occurred during the preshift period
<i>short-term training</i>	training that occurs for a short term
<i>downslope movement</i>	something moves down slope
<i>the chamber concentration</i>	the concentration of something in the chamber

Finally, there are many other combinations that express more idiosyncratic semantic relations to a nominalized head noun. Thus consider:

<i>unit variance</i>	something varies with a unit of one
<i>the Dred Scott decision</i>	the courts decided something in the case involving Dred Scott
<i>a two-stage decision</i>	someone decided something, and the process required two stages
<i>heat treatment</i>	someone treated something, and heat was the instrument
<i>climate information</i>	someone informed someone about climate

<i>hand calculations</i>	someone calculated something, using 'by hand' methods
<i>human-environment interaction</i>	humans interact with the environment
<i>Monte Carlo variance</i>	something varies in a certain way determined through Monte Carlo methods of analysis

As illustrated by the glosses for these combinations, the meaning relationships between the two nouns in these constructions can be made explicit through the use of relative clauses or other forms of clausal modification – the preferred form of expression in earlier centuries. In contrast, the noun–noun sequences presented above, which are typical of late-twentieth century research writing, are completely inexplicit regarding the intended meaning relationship between the two nouns.

In most cases, the individual nouns used as NP pre-modifiers are not rare or technical words (e.g., nouns like *hand*, *heat*, *climate*, *two-stage*, *unit*, *spring*, *government*, *drug*, and *teacher* in the aforementioned examples). However, because the meaning relationship between the modifying noun and head noun can be so varied, with no explicit indication of the relationship, these combinations can be extremely difficult to comprehend, especially for non-expert readers.

For example, the following sentence, taken from a psychology research article, contains three noun phrases with a noun pre-modifier:

In intermittent punishment training, then, the sign function of nonpunishment would not be one of anxiety reduction.

The last noun phrase with a noun pre-modifier, which actually utilizes the most technical vocabulary, is probably the easiest to understand: 'anxiety' is the semantic patient of 'reduction' (i.e., something reduces anxiety). However, the other two cases are much more difficult to understand, even though the pre-modifying noun is not especially technical. A careful reading of this article shows that *punishment training* refers to a kind of training where subjects are punished for incorrect behavior. Thus, 'punishment' is neither the agent nor patient of 'training'; it rather has a kind of specialized instrumental semantic relationship to 'training'. The interpretation of *sign function* is even more problematic: this combination occurred only once in this research article, and the term 'sign' was also not used again in the article. We thus have no idea what the intended meaning is for this combination. Such combinations are not at all uncommon in specialist academic prose; the individual words are usually easy to understand, but without specialist

background knowledge, it is not possible to determine the exact intended meaning relationship between the modifier and the head noun.

We have discussed nouns as noun pre-modifiers in detail because they are the most frequent compressed structural device. However, all phrasal forms of modification preferred in academic writing are inexplicit in similar ways. For example, nouns are often compounded with adjectives, which are in turn used to pre-modify a head noun. These structures incorporate multiple complex but inexplicit meaning relations, as in:

aspirin-resistant patients

[compare: patients who resist aspirin]

an ATP-dependent conformational change

[compare: a change that is conformational and that depends on ATP]

Noun-participle compounds functioning as NP pre-modifiers are similarly inexplicit. The use of these structures enables an extremely dense packaging of information, as in:

This study considers the role of the cytoplasmic segments of Sed61a in the ribosome-binding and translocation-promoting activities of the Sec61 complex.

actinomycin-treated, RNA-depleted cells effectively support the growth of Mengovirus meanings

Although there are no overt grammatical signals, the grammatical meaning relations underlying 'N1-*ing*-participle N2' constructions are relatively easy to process, because they always express a single meaning: the pre-modifying noun (N1) is the semantic patient of the *-ing* verb, and the head noun (N2) is the effective agent of the *-ing* verb. Thus, compare:

reef-building corals

[compare: the corals build the reefs]

electron-scattering material

[compare: the material scatters the electrons]

ribosome-binding and translocation-promoting activities of the Sec61 complex

[compare: the activities of the Sec61 complex bind ribosome and they promote translocation]

In contrast, 'N1-*ed*-participle N2' constructions can represent a range of different meaning relations. Some of these constructions can be directly paraphrased as a non-finite passive voice post-modifying clause, as in:

a neutron-activated specimen

[compare: a specimen activated by a neutron]

a lead-shielded crystal

[compare: a crystal shielded by lead]

wind-baked ruffians

[compare: ruffians baked by the wind]

a Puritan-ruled state

[compare: a state ruled by Puritans]

teacher-requested prosocial behavior

[compare: prosocial behavior requested by the teacher]

But in many other cases, the compounded *-ed* participle needs to be expressed as a prepositional verb when the structure is rephrased. The specific preposition has been omitted in the compounded pre-modifying structure, but it is required in the clausal construction. For example:

research-based instructional improvement

[compare: instructional improvement based on research]

sex-related differences

[compare: differences related to sex]

actin-linked mechanisms

[compare: mechanisms linked to actin]

goal-directed play

[compare: play directed towards a goal]

infusion-treated patients

[compare: patients treated with an infusion]

In some cases, the omission of the required preposition in the pre-modifying compound can actually lead to uncertainty regarding the intended meaning. For example, consider the following example, containing two 'N1-*ed*-participle N2' modifiers:

The granules are best seen in sections of OSO₄-fixed, plastic-embedded tissue.

These pre-modifying compounds might be paraphrased with different prepositions:

tissue fixed by OSO₄ versus tissue fixed with OSO₄

tissue embedded in plastic versus tissue embedded with plastic

And this choice of preposition can result in fundamentally different meanings; for example:

tissue embedded in plastic [compare: the tissue is embedded in the plastic]

versus

tissue embedded with plastic [compare: the plastic is embedded in the tissue]

Thus, for the non-specialist reader, the inexplicit meaning relationships inherent with many attested examples of these 'N₁-*ed*-participle N₂' constructions may lead to lack of comprehension, ambiguity of meaning, or even incorrect interpretations of the relationship between N₁ and N₂.

More detailed consideration of 'N₁-*ed*-participle N₂' constructions shows that they are inexplicit at an even more fundamental level, relating to the semantic relationship between the two nouns and the participle. In cases where the participle can be paraphrased as a simple verb, the pre-modifying noun (N₁) is the semantic agent of the *-ed* verb, and the head noun (N₂) is the patient of the *-ed* verb. For example:

a neutron-activated specimen

[compare: a specimen activated by a neutron >> a neutron activates the specimen]

a lead-shielded crystal

[compare: a crystal shielded by lead >> lead shields the crystal]

wind-baked ruffians

[compare: ruffians baked by the wind >> wind baked the ruffians]

a Puritan-ruled state

[compare: a state ruled by Puritans >> Puritans ruled the state]

teacher-requested prosocial behavior

[compare: prosocial behavior requested by the teacher >> the teacher requested the prosocial behavior]

However, cases where the participle would be paraphrased as a prepositional verb represent a completely different semantic relationship: *both* nouns are complements of the verb, and the agent is unspecified. For example:

research-based instructional improvement

[compare: instructional improvement based on research >> someone based instructional improvements on their research]

sex-related differences

[compare: differences related to sex >> someone related those differences to sex]

actin-linked mechanisms

[compare: mechanisms linked to actin >> someone linked related those mechanisms to actin]

goal-directed play

[compare: play directed towards a goal >> some group of children directed their play towards a goal]

infusion-treated patients

[compare: patients treated with an infusion >> some medical professional treated patients with an infusion]

The inexplicitness of these constructions is even more complicated when there is another pre-modifying noun, resulting in the 'N1-*ed*-participle N2 N3' pattern. In some cases, it is relatively easy to infer that the N1-*ed*-participle compound modifies N2, but in other cases, we can infer that the N1-*ed* compound modifies N3. Thus contrast:

a) Accountability-based school reform is not without paradox.

versus

b) This organization is completely student run and operates as a collective of five student-initiated recruitment centers.

In (a), it is relatively easy to infer that the 'reforms' (N3) are based on 'accountability' (N1); it would not make sense for the 'schools' (N2) to be based on 'accountability' (N1). However, in (b), the opposite situation holds, where we can infer that the 'recruitment' (N2) is initiated by 'students' (N1); in this case, it makes little sense for the 'centers' (N3) to be initiated by 'students' (N1).

In some cases, inferences like these are facilitated by the core meanings of the nouns (N2 and N3). However, in many other cases, the N1-*ed*-participle compound could easily modify either noun, and expert background knowledge is required to determine the intended meaning. This is the case even when the specific nouns are non-technical and familiar to the reader. For example, contrast the following:

The present study was based on [...] a decade of accountability-focused reform efforts.

[compare: the reforms (N2) are focused on accountability

versus

the efforts (N3) are focused on accountability]

A range of other legislative and court-based change efforts had been attempted.

[compare: the change (N2) is based on courts
versus
the efforts (N3) are based on courts]

NCES-constructed student-level panel weights [...] can be used to generate national population estimates.

[compare: the panels (N2) are constructed by the NCES
versus
the weights (N3) are constructed by the NCES]

The prevalence of airway obstruction and self-reported disease status were calculated. . .

[compare: the disease (N2) was reported by self
versus
the status (N3) of the disease was reported by self]

Of course, the inexplicitness of intended meaning becomes even more problematic with longer structures or with structures that employ technical vocabulary. Thus, consider the following examples:

Moreover, yeast centromeres cluster near the membrane-embedded spindle pole body.

With the combination of test-generated error variance and between-cohort variance, it was inevitable that schools would produce a certain amount of 'noise'.

. . . on the mechanisms of ATP-induced shape changes in human erythrocyte membranes.

. . . predicting the impacts of harvesting using stage-structured population models

An ARU:550 indicates the absence of aspirin-induced platelet dysfunction.

. . . the multiple actions of GLP-1 on the process of glucose-stimulated insulin secretion.

In all the aforementioned examples, the compound is hyphenated (e.g., *accountability-focused*, *self-reported*). In this case, the hyphen provides an overt signal that the participle should be interpreted as the modifier of the preceding noun (rather than as a modifier of the following noun). However, it has recently become common in science writing for these

structures to occur without the hyphen – a shift towards an even less explicit representation of the meaning relationship among elements. It turns out that these sequences can in fact be used to express different meaning relations. In some cases, these sequences function as if there were a hyphen, so that the *-ing* participle modifies N₁, and the entire N₁ + *-ing* participle compound modifies the head noun (N₂). In these structures, N₁ functions as the semantic patient of the *-ing* participle. For example:

PLK1 treatment partly inhibited the filament forming ability of vimentin

[compare: the ability to form filament, but NOT abilities for forming and abilities for filaments]

... each experiment was initiated with 1.1×10^6 colony forming units

[compare: units that form colonies, but NOT units that form, and units for colonies]

However, there are other structurally-similar examples where both N₁ and the *-ing* participle directly modify the head noun (N₂); for example:

weapons training program

[compare: a program that trains people in the use of weapons;
that is, a program for training, which is also a program for weapons, but
NOT a program that trains weapons]

high resolution imaging techniques

[compare: techniques for imaging, which are also techniques that pro-
duce high resolution, but NOT techniques for imaging high resolution]

The aforementioned examples are difficult to comprehend, even when they employ non-technical vocabulary. Longer sequences become even more difficult to process. Thus, consider:

item difficulty rating screens

[compare: screens where participants can rate the difficulty of items]

item difficulty rating accuracy

[compare: the accuracy with which participants have rated the difficulty
of items]

ground truth data sets

[compare: sets of data that represent truthful information as it is
observed on the ground]

Furthermore, as technical vocabulary is utilized, the inexplicitness resulting from these compressed phrasal constructions becomes even more apparent to the non-specialist reader. In such cases, a non-specialist reader may have little knowledge of the possible meaning relationships between the head nouns and their modifiers. The lack of knowledge about the content meanings of the individual words makes it even more difficult to form hypotheses about the intended meaning of the construction;¹ for example:

mRNA decapping trigger apoptosis in yeast

GPI-mediated PRP membrane attachment

a high-accuracy human primary T cell signaling causality map

6.4 Inexplicit meaning relationships associated with phrasal post-modifiers in noun phrases

Most NP post-modifiers have an overt grammatical signal that marks the existence of an embedded structure, providing an indication of the meaning relationship between the head noun and modifier. Finite relative clauses provide the most grammatical information: a full verb phrase that expresses tense, aspect voice, and modality, together with the main verb itself, which expresses a specific meaning relationship between the head noun and the other constituents of the relative clause. Non-finite relative clauses omit the expression of tense/modality, but they are otherwise as explicit as finite relatives. Prepositional phrases as noun modifiers are considerably less explicit, because the preposition is the only indication of the meaning relationship between the head noun and modifying noun phrase (see the following discussion). However, at the opposite extreme from finite relative clauses, appositive noun phrases are similar to NP pre-modifiers in that they have no overt grammatical signals to indicate the intended meaning relation to the head noun.

As described in earlier chapters, appositive structures comprise two noun phrases that occur in sequence, separated by a comma or

¹ The recent book by Feist (2012) provides a description of nominal pre-modifiers in English focused mostly on the semantic principles governing the order of pre-modifiers. These principles provide guidance to the semantic interpretation of pre-modifiers used in different semantic 'zones' (epithet, descriptor, classifier), but the treatment has less to offer regarding the wide range of specific meaning relationships among these elements. Future research is likely to identify additional regularities (perhaps associated with different sets of head nouns and modifiers) that help to predict the meaning relationships found in particular noun phrase constructions. However, based on our survey to date, we expect that many of these are simply inexplicit in meaning, requiring the background knowledge of expert specialist readers to determine the intended relationships among elements.

parentheses. Traditionally, the meaning relationship is straightforward, and thus there was really no need for overt grammatical signals: the two noun phrases are simply co-referential, with each one providing different descriptive information. The typical use in newspaper writing is even more constrained, where the first noun phrase is often a proper noun, and the second noun phrase clarifies the identity or provides descriptive information that would help to specify the reference. The examples follow this pattern, with NP₁ (**bolded**) as a proper noun and the appositive NP₂ (underlined) clarifying the identity of that proper noun:

*And so in the middle of **Chuseok**, the Korean Thanksgiving, Ban was on the phone to his counterparts in Moscow.*

*"This will be the first time he's ever been his own boss," says **Peter Beck**, the Seoul-based director of the International Crisis Group's Northeast Asia project.*

In such cases, the meaning relationship is not problematic, because it is almost always a simple identity; for example, *Chuseok* **is** *the Korean Thanksgiving*, and *Peter Beck* **is** *the Seoul-based director*.

In some cases, multiple appositive noun phrases can be embedded, with each one providing additional information about the referent. For example, the following sentence contains two appositive noun phrases, with the second appositive modifying a head noun within the first appositive:

*[The military has made] **six daily runs to deliver supplies to some of Indonesia's most remote villages** – all part of **Operation Unified Assistance**, the largest U.S. military operation in Asia since Vietnam and the backbone of the global campaign to fend off hunger.*

In this case, the first NP '*six daily runs...*' **is** *a part of Operation Unified Assistance*, and that noun phrase **is** *the largest U.S. military operation...*

However, the use of appositive noun phrases is often less predictable, and therefore less explicit, in academic writing. In some cases, appositive noun phrases are used generally with identity meaning relationships, but even these represent many different specific functions. For example, in humanities academic writing, appositives are often used to identify a person, in a similar way to newspaper writing:

*More fundamentally, what role had **Montagu Norman**, the Bank's governor, and **Sir Otto Niemeyer**, the director in charge of its overseas and foreign department, played...*

*The Nazis attempted to secure this gold by threatening **Malik and Peroutka**, the relevant NBC directors.*

More often, though, appositives in academic writing are used to modify an indefinite head noun phrase. In these cases, knowing the specific identity of the head noun is not crucial for the main argument, but providing the specific identity adds credibility. This use is common in both humanities and science writing; for example:

*Plots against Queen Elizabeth revealed that the first allegiance of Catholics was to **a foreign prince** (the pope).*

Two authors (CD, KS) independently abstracted data from all studies using standardized forms.

*Disagreements between the primary abstracters were resolved by **a third author** (SS).*

***Two Stuart monarchs** (Charles I and Charles II) were strongly suspected of Romish sympathies*

*Early-onset pneumonia is usually caused by microaspiration of **bacteria colonizing the oropharynx** (gram-positive cocci and Haemophilus influenza)*

*In a field study, we examined a guild of six **large-bodied, lotic insects** (Plecoptera and Trichoptera)*

In other cases, the appositive provides a definition for a technical term, or vice versa:

*These are the three faculties of the soul that, according to the rhetorical tradition, make eloquence possible and need to be developed through exercise: **memoria** (the ability to remember one's speech as well as the opponent's arguments), **ingenium** (natural intelligence and imagination), and **judicium** (good sense and good taste).*

*In matters of **literary exegesis** (enarratio poetarum) considering decorum means interpreting a text with regard to its historical context.*

*A case study of **wild Atlantic salmon parr** (salmo salar) showed. . .*

A related function is to provide an acronym for a technical phrase, as in:

*The risk of this complication in the **intensive care unit** (ICU) ranges from 8% to 25%.*

***Length at first capture** (LFC) adjusts for a spurious effect resulting from a data bias.*

*Pulmonary involvement was defined as bibasilar pulmonary fibrosis as defined above; or crackles ("velcro" rales) plus **forced vital capacity** (FVC). . .*

Although these specialized functions differ from the typical uses in newspaper writing, they can all be regarded as sub-types of the same general function of expressing an identity relationship between two noun phrases.

However, it turns out that these are only a small subset of the full range of meanings expressed by appositives in academic research writing. For example, in humanities prose, appositives can be used to provide background commentary about an event or concept referred to by the head noun:

*The civil war ended with **the triumph of the Bolshevik regime in Russia**, a development that had been increasingly obvious for some months prior to late 1920.*

*Blackstone often explained away apparent absurdities and contradictions in **his beloved common law**, the source and guarantee of English liberty, as vestiges of "the blind days of popery."*

In science research writing, appositives are often used to provide methodological details, describing the way in which a construct was measured or operationally defined in a study. In some cases, appositives identify the role of a variable in the research design, as in:

*Studies were included if mechanically ventilated patients were prospectively assigned randomly to some form of subglottic secretion drainage versus no **drainage** (control).*

***Rehearsal technique** (the independent variable) had five levels.*

In other cases, appositives simply identify the technical procedures used for a measurement, as in:

*In this study we use **measurements of body size** (fork-length and wet-weight) to derive empirical models.*

This latter function is quite common in science and medical research writing, where it is important to specify how variables were measured for the purposes of the study. Appositives are commonly used to provide these methodological details in a structurally compressed manner; for example:

*Initial scleroderma-related changes included proximal scleroderma; sclerodactyly; **puffy hands** (sausage-like edema lasting > 6 months). . .*

*We also applied the American Society of Echocardiography recommendations, in which the diagnosis of **pseudonormal filling pattern** (early/late diastolic*

*flow: 1.0 to 2.0) and **restrictive filling pattern** (early/late diastolic flow > 2) requires **evidence of structural heart disease** (left atrial enlargement, left ventricular enlargement or hypertrophy).*

***Trace quantities of algae or detritus** (the amount that might be found in gut contents of their prey) were assigned a nominal value of 0.1.*

*Hunters registered a total number of 305 roaring males during 4161 **observation hours** (observation time × number of observers).*

*Applying the scale in a nonblinded fashion to the studies in our meta-analysis reveals that they are all of **moderate quality** (scores of 2 to 4 on a scale of 0 to 5).*

In addition, appositives are commonly used to present research findings, such as the frequency for an observed pattern, or the specific results of a statistical analysis:

*Causes of death were **renal failure** ($n = 30$), **cardiac failure** ($n = 42$), **pulmonary involvement** ($n = 35$), and **pulmonary arterial hypertension** ($n = 41$).*

*Involvement during follow-up was also related to **the presence of anti-topoisomerase I antibodies** ($r = 0.20$; $P < 0.001$).*

*Secretion drainage shortened **the duration of mechanical ventilation** by 2 days (95% CI: 1.7 to 2.3 days) and **the length of stay in the intensive care unit** by 3 days (95% CI: 2.1 to 3.9 days), and delayed **the onset of pneumonia** by 6.8 days (95% CI: 5.5 to 8.1 days).*

Similarly, appositives are commonly used to identify a table or figure that presents the specific results corresponding to a generalized pattern described by the head noun phrase:

*All cohorts showed **significantly increased standardized mortality ratios** (Figure 1).*

*This study demonstrates large uncertainties in predicting the future growth of an introduced small population of muskox in Central Norway, owing mainly to **the effects of environmental stochasticity** (Fig. 1).*

*The Medline search retrieved 110 citations, of which five met **our inclusion criteria** (Table 1).*

In many cases, scientific prose employs an extremely high density of appositives, even with embedded appositives, each representing a different

specific meaning relationship to its head noun. Thus, consider the following examples:

*The individual cohorts comprising the database differed significantly ($P < 0.01$ for all) in **mean age at enrollment** (range, 43.7 [Keio] to 55.4 [Mayo] years); **proportion of men** (5% [Mayo] to 39% [Nijmegen]); and **diffuse cutaneous involvement at enrollment** (8% [Mayo] to 51% [Pittsburgh]).*

*The use of subglottic secretion drainage reduced **the risk of ventilator-associated pneumonia by about half** (summary risk ratio [RR] = 0.51; 95% confidence interval [CI]: 0.37 to 0.71). There was a similar reduction in **the incidence of ventilator-associated pneumonia** (summary rate ratio = 0.57; 95% CI: 0.33 to 0.97). [. . .] The effects of subglottic secretion drainage on several outcomes were significantly different among the studies summarized, specifically for the **duration of mechanical ventilation** (heterogeneity $P = 0.004$), **length of ICU stay** (heterogeneity $P < 0.001$), and **time to onset of pneumonia** (heterogeneity $P < 0.001$) (Figures 2 and 3).*

*We compared **breeding phenology**, i.e. **the timing and synchrony of rutting** (roaring, sexual aggregation) and calving of **red deer** (*Cervus elaphus* L.) in **France** (latitude: 49°N) and **Norway** (latitude: 63°N).*

In summary, appositive noun phrases can be used to represent an extremely wide range of meanings in relation to the head noun. Like many other phrasal devices used to modify head nouns in modern academic writing, there are no overt grammatical signals of those meaning relationships in the appositive structure. As a result, these features provide yet another important example of the inexplicit packaging of information in academic research writing.

Prepositional phrases as NP post-modifiers differ from other phrasal noun-modifying devices in that they have a grammatical signal of the meaning relationship between the modifier and the head noun: the preposition (e.g., *in*, *on*, *for*, *with*). On first consideration, it might appear that these prepositions provide a clear, unambiguous expression of the meaning relationship. For example, *in* means inside; *on* means a position on top of something; *beside* means next to something; and so on. In academic writing, though, these prepositions rarely express such explicit locative (or concrete) meanings. Rather, the same preposition can be used to express a wide range of different, and often abstract, meaning relationships.

For example, consider the three prepositional phrases beginning with *for* in the following sentence:

*Another reason to use Ohio as a **surrogate** [for the country as a whole] is that the **data base** [for hazardous waste generation and flow] [for the State] is fairly good.*

In their grammatical form, these three phrases are identical. Yet, they express three different meaning relationships. We can see those differences by considering a clausal paraphrase, which would make the meaning relationships explicit:

*a **surrogate** [for the country as a whole] >>*

a surrogate that represents the country as a whole

*the **data base** [for hazardous waste generation] >>*

the data base that documents hazardous waste generation

*the **data base** ... [for the State] >>*

the data base that the State uses

Other meaning interpretations could have been possible here, at least for the non-expert reader. For example, *the data base for hazardous waste generation* could mean 'the data base used to generate hazardous waste'. Similarly, *hazardous waste generation and flow for the State* could mean 'the hazardous waste that is generated and flows in the State' or 'the hazardous waste generation and flow that is produced for the purposes of the State'. Thus, although the preposition *for* restricts the range of possible meanings, it provides no explicit expression of the specific intended meaning. Rather, meaning is implicit and must be inferred by the reader.

Prepositional phrases with other prepositions are similarly inexplicit. For example, consider the wide range of meaning relationships associated with prepositional phrases beginning with *in* and *on*:

farms in Malaysia

[compare: farms that are located in Malaysia]

experiments in India

[compare: experiments that were conducted in India]

experiments in agricultural chemistry

[compare: experiments that focused on the study of agricultural chemistry]

a decrease in oil prices

[compare: oil prices decreased]

the roots on the surface

[compare: roots which are on top of the surface]

restrictions on underground injection of chemicals

[compare: rules that restrict the underground injection of chemicals]

writers on style

[compare: writers who discuss style]

The loss of meaning associated with noun-modifying prepositional phrases becomes especially apparent when we attempt to read technical prose about specialized topics. For example, consider the phrase *reflection of Alfven waves* in the following sentence:

Interest is now developing in a theoretical approach involving reflection [of Alfven waves.]

For the non-expert reader, two meanings are possible:

'a theoretical approach which involves the way in which Alfven waves reflect something'

or

'a theoretical approach which involves the way in which something reflects Alfven waves'

Similar potential for misunderstanding by novice readers arises when there are multiple prepositional phrases in sequence. In this case, readers must determine which head noun is being modified by each prepositional phrase, in addition to determining the intended meaning relationship represented by the preposition. For example, consider the following sentences from an ecology article about food web structure:

Therefore, in our case it was possible to find an association between the number of trophic connections and productivity for individual predator species.

[compare: 'an association that holds for (= in the case of) individual predator species' or 'productivity for (= associated with) individual predator species']

Time lags suggest the involvement of population processes in determining observed patterns.

[compare: 'involvement in determining' or 'processes in determining']

Among predator populations, changes in abundance often imply variation in the proportions of species with different feeding strategies.

[compare: 'the proportions with different feeding strategies' or 'species with different feeding strategies']

6.5 Inexplicit expression associated with compressed clausal connectors in academic writing

All of the grammatical devices discussed in the previous sections are phrasal modifiers of head nouns. They are inexplicit in that they omit grammatical markers that could specify the meaning relationship between the head noun and the modifier. However, there is also evidence that this style of discourse is spreading to a less explicit marking of meaning relations generally in academic texts. One reflection of this general trend is a shift in the way that authors mark the meaning relationships among independent clauses. Traditionally, authors employed linking adverbials for this purpose. These are clause connectors like *however*, *therefore*, *thus*, *for example*, *furthermore*, and *in addition*. Linking adverbials have a peripheral syntactic role in their matrix clause; their primary function is to specify the logical relationship of a clause to other clauses in the discourse. Thus, consider the role of linking adverbials (**bold underlined**) in the following passage from an education research article:

Since the OPI items have not been rated for social desirability, the study has no direct bearing on this related body of research. **However**, it is probable that when subjects are instructed to make a good or a bad impression, the perceived social desirability of the items becomes relevant. **Thus**, those scales that are most susceptible to faking are probably those composed of items whose content reflects the most socially desirable and undesirable attitudes and opinions.

From a syntactic perspective, linking adverbials like *however* and *thus* are optional: the grammatical structure of the matrix clause is not affected by the presence or absence of a linking adverbial. However, these clause connectors serve an important discourse function in explicitly identifying the logical relations among independent clauses. Thus, in the aforementioned example,

Table 6.1. Statistical comparisons for clause connectors signaling logical relations

Linguistic feature	ANOVA statistics	Mean Scores (per 1,000 words)			
		Humanities	Social science	Popular science	Specialist science
linking adverbials	$F = 26.0$; $p < .0001$; $r^2 = .28$	7.4	7.9	8.4	4.4
colons	$F = 39.6$; $p < .0001$; $r^2 = .37$	2.4	4.7	1.3	4.2

the use of *however* tells the reader that the following independent clause will contradict the assumption of ‘no direct bearing’ expressed in the preceding clause. And the use of *thus* in the final clause signals that the following information will explicate the statement made in the preceding clause.

Linking adverbials – similar to finite dependent clauses – are most common in humanities and social sciences academic prose, but considerably less common in specialist science research articles. In fact, Table 6.1 shows that there are only about half as many linking adverbials in specialist science articles as in the other academic sub-registers.

Table 6.1 also presents the results for colons across academic sub-registers. Although colons have no semantic content, they can be regarded as an alternative device used to connect independent clauses.² As Table 6.1 shows, colons have a quite different distribution across disciplines from linking adverbials: they are common in social science and specialist science research writing, but much less common in humanities and popular science writing.

Similar to the other grammatical devices that are especially characteristic of present-day science writing, there has been major historical change in the use of these clause-connecting devices. Figure 6.1 shows that linking adverbials were quite common in academic writing in the eighteenth and nineteenth centuries, but they have decreased rapidly in use in the twentieth century. Linking adverbials that overtly express meanings of logical consequence or contrast (*therefore*, *thus*, and *however*) are less than half as frequent in present-day research writing as they were in 1800. In contrast, colons have rapidly increased in use, especially during the twentieth century.

² The semi-colon has a more complicated history, because it can be used to connect either clauses or phrases.

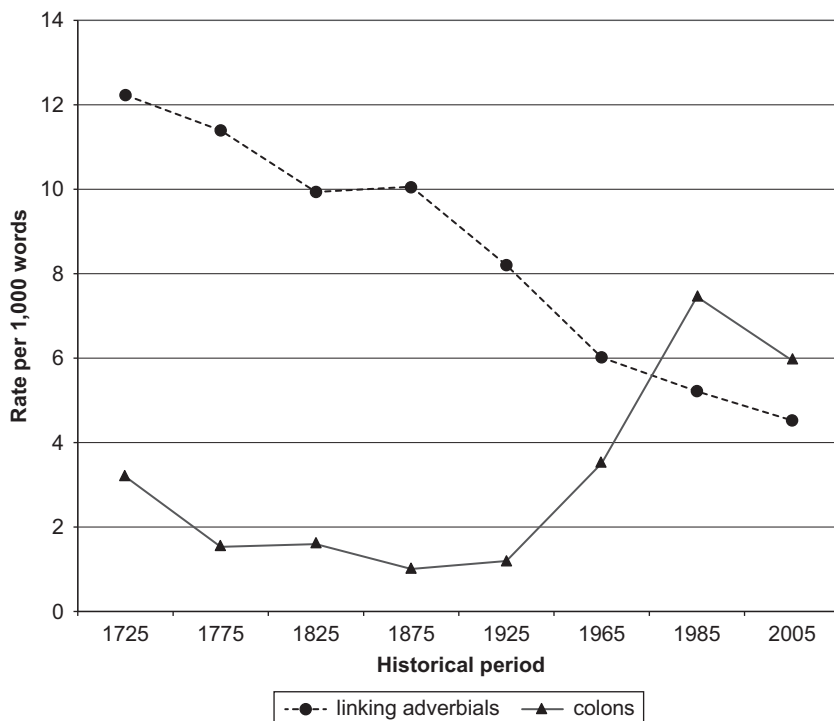


Figure 6.1 Historical change in the use of linking adverbials versus colons as clause connectors in academic prose

Here again we see the general historical trend (especially in science research writing) away from the explicit marking of meaning relations towards a discourse style that is more compressed structurally, and as a consequence, less explicit in meaning. That is, similar to the other grammatical devices discussed earlier, the meaning relation signaled by a colon is entirely implicit.

In academic writing, there are several different meaning relations that can be signaled by a colon. In some cases, a colon is used as an alternative to a comma to connect two appositive noun phrases, as in:

*No understanding of slavery can avoid **these themes**: violence, power, and the usurpation of labor for the purpose of aggrandizing a small minority.*

The colon is also often used to connect appositive noun phrases at a distance. For example:

One possible pathogenetic mechanism can be excluded with certainty in man:
a physiological decrease of intestinal lactase activity with advancing age.

In addition, the colon can be used for clausal rephrasing, with the second clause elaborating or explaining the content of the first clause:

This is indicated in Figure 2b by the difference in slopes of the mean correct-response curves: The higher 8-minute recovery curve rises more steeply than the 1-minute curve.

However, clauses connected by a colon can also have an explanatory or causative relationship. Thus, the following sentence could be paraphrased with a reason clause:

If replication of the viral nucleic acid is dependent, there is no need for special enzymes in order to carry it out: the normal cellular enzymes should be sufficient.
[compare: there is no need ... because the normal cellular enzymes should be sufficient]

In all of these cases, there is no overt grammatical device to specify the meaning relationships. The colon signals only that some meaning relationship exists between the following phrase/clause, and some preceding phrase/clause. Thus, similar to phrasal noun-modifiers, colons are structurally compressed, resulting in more compact discourse, which in turn results in less explicit signaling of the meaning relationships in specialist academic discourse.

In this chapter, we have highlighted many of the ways in which the compressed, phrasal discourse style of academic writing results in inexplicit meaning relationships. These inexplicit meanings are in fact contrary to the often-held perceptions of academic writing as explicitly encoding logical relationships. In [Chapter 7](#), we turn to a discussion of the implications of this compressed, inexplicit discourse style.

Conclusion

We have addressed three major topics in the present book, identified in the book's title: grammatical complexity, written academic English, and historical change. In our early thinking about the book, we had planned to do a study of linguistic change in the use of grammatical complexity features. We were especially interested in two major patterns that had been generally disregarded in previous research: the way in which grammatical complexity is realized through phrasal devices rather than through dependent clauses, and the way in which historical change in the use of those features has emerged and evolved in written discourse, rather than in spoken interaction.

In carrying out our research on these two patterns, we consistently came to the conclusion that academic writing – and especially the written discourse of specialist science research articles – is one of the most important registers in English for the study of these historical developments. And thus, we added a third major objective for the book: to provide a relatively comprehensive account of the distinctive grammatical style of academic research writing, contrasting the different types of complexity found in spoken versus written registers, and focusing on the distinctive complexity features that are especially typical of academic science writing.

In part, our motivation for these three objectives stemmed from several major stereotypes prevalent in previous research about grammatical complexity, the process of language change, and academic writing. We identified these major stereotypes in [Chapter 1](#), and then challenged them throughout the chapters in this book by presenting empirical documentation of the linguistic patterns of use in academic writing, contrasted over time and across registers. In the following sections, we revisit these stereotypes and summarize how the evidence presented in the book has problematized each one, before turning to the more applied implications of our findings.

7.1 Stereotypes about grammatical complexity: What does it mean to be ‘complex’?

Traditionally, linguists have operationalized grammatical complexity as the increased use of ‘elaborated’ structures, consisting of dependent clauses added on to a simple independent clause. That is, a sentence is considered complex to the extent that it has multiple embedded subordinate clauses.

We do not dispute the claim that ‘elaborated’ clauses with embedded dependent clauses represent one type of grammatical complexity. However, the central themes in our book include the complementary claims that:

- there are different types of grammatical complexity;
- one of these alternative types of complexity involves the extensive use of embedded phrasal structures (rather than embedded clauses);
- phrasal complexity features are much more characteristic of, and thus more important in, academic writing than clausal complexity features; and
- these phrasal complexity features have been the locus of historical change in academic registers over the last three centuries.

We have amply illustrated how phrasal embedding – especially phrasal features functioning as noun phrase modifiers – can result in sentences that are grammatically complex. We have further shown that these kinds of complexity are less explicit, and therefore more complex from a processing perspective, than alternative structures with dependent clauses. And we have shown that these phrasal complexity features are highly specialized: they are relatively recent innovations, and they occur primarily in academic (science) research writing. That is, these phrasal complexity features are ubiquitous in that register, resulting in discourse passages like the following:

A basis for the interpretation of the structure of the cell membrane is often looked for in electron microscope investigations on the structure of lipid models. A difficulty in these investigations is our lack of knowledge of the effect of the preparative treatment on the structure studied. This applies especially to the strongly oxidizing fixatives: osmium tetroxide and potassium permanganate.

P.F. Elbers, P.H. Ververgaert, and R. Demel. 1965.
‘Tricomplex fixation of phospholipids’
Journal of Cell Biology, 2: 23–30.

Each of the sentences in this passage has a single main verb (*is often looked, is, applies*), and only one of these sentences has a dependent clause (with the non-finite verb *studied*). If we restricted our definition of complexity to consider only the use of dependent clauses, we would conclude that this passage was not complex – a conclusion completely at odds with our perceptions as readers! This apparent contradiction reflects the fact that the sentences in this passage are in fact very complex in their reliance on phrasal modifiers of noun phrases, including:

- pre-modifying nouns (e.g., *cell membrane, electron microscope investigations*),
- attributive adjectives (e.g., *preparative treatment*),
- prepositional phrases as post-nominal modifiers (e.g., *a basis for the interpretation..., a difficulty in these investigations...*),
- appositive noun phrases (e.g., *the strongly oxidizing fixatives: osmium tetroxide and potassium permanganate*).

One of our main goals in the book has been to advocate this alternative perspective on grammatical complexity, showing that embedded phrasal structures are at least as important as embedded dependent clauses for characterizing grammatical complexity. In addition, we have shown that grammatical complexity is *not* a single unified construct. Rather, there are multiple types of grammatical complexity, and these types differ in important ways with respect to their structural and syntactic characteristics, their discourse functions, and their distribution across registers. Accordingly, any attempt to analyze grammatical complexity through a single omnibus measure is misguided. Quantitative measures such as sentence length or t-unit length – motivated by the belief that there is a unified construct of grammatical complexity – fail to capture the important differences associated with clausal versus phrasal complexity features. This recognition is important both theoretically and practically. We return to the applied implications of these research findings in [Section 7.4](#).

7.2 Stereotypes about historical change: Does all grammatical change originate in speech?

In [Chapter 1](#), we documented how linguists have often staked out an extreme position on the status of written discourse: speech is often described as ‘real language’, while written discourse is often dismissed as

a secondary artifact. When it comes to the study of historical change, linguists complain that we are stuck with written texts as our only indication of language use in earlier centuries. This is often portrayed as an unfortunate situation, because direct analysis of written discourse is claimed to have little theoretical interest in itself. Rather, the assumed task of the historical linguist is to remove the obscuring filter of the written form, to discover historical change as it really occurred in speech, which is assumed to be the real language underlying written discourse. Some linguists acknowledge that written discourse also changes, but the generally accepted view is that speech takes the lead in linguistic innovations, with written discourse trailing along behind: ‘the spoken form always, always keeps on changing—and ultimately drags the written form reluctantly with it’ (McWhorter 2001: 17). Academic writing is regarded as especially conservative and resistant to historical change: one of the ‘up-tight’ registers identified by Hundt and Mair (1999).

In the present book, we have challenged these stereotypes at every level. We adopt the ‘register’ perspective on language use (see Biber and Conrad 2009), which begins with an analysis of the situational context of a text, followed by a linguistic analysis of the distinctive grammatical characteristics of the text. The patterns of linguistic variation among texts are then interpreted in functional terms, based on the claim that:

linguistic features are always **functional** when considered from a register perspective. That is, linguistic features tend to occur in a register because they are particularly well suited to the purposes and situational context of the register. (Biber and Conrad 2009: 6)

From this perspective, all registers (spoken or written) will have distinctive linguistic characteristics, which are functionally associated with their situational context of use.

The register perspective can be considered as an extension of earlier research carried out within the general framework of ‘functional linguistics’:

[Functional grammar] analyzes grammatical structure [. . .] but it also analyzes the entire communicative situation: the purpose of the speech event, its participants, its discourse context. Functionalists maintain that the communicative situation motivates, constrains, explains, or otherwise determines grammatical structure [. . .] Functional grammar, then, differs from formal and structural grammar in that it purports not to model but to explain; and the explanation is grounded in the communicative situation. (Nichols 1984: 97)

Research carried out within the framework of ‘functional linguistics’ has usually been restricted to language use in spoken interaction. In contrast, the perspective of register variation adopted in the present study is based on the premise that *every* register (written as well as spoken) has a situational context and communicative purpose that functionally ‘motivates, constrains, explains, or otherwise determines grammatical structure’ (Nichols 1984: 97). It is only logical that these functional forces operate diachronically as well as synchronically. As a result, far from being a filter that obscures language change, we have argued that written discourse provides the context needed for certain types of linguistic change – that is, that certain types of linguistic change occur in writing due to the situational context of written language. In particular, we have shown how the informational purposes, careful production circumstances, and specialist audiences of scientific research writing are functionally associated with dramatic historical innovations and functional extensions for a range of phrasal complexity features.

7.3 Stereotypes about academic writing: Is it resistant to change, elaborated, and explicit?

It turns out that academic writing is especially distinctive relative to the range of other registers: distinctive in the types of grammatical complexity employed in texts, and distinctive in the patterns of historical change. And so the present book has focused on the grammar of academic writing – and as a result, challenged multiple stereotypes about that register.

For example, linguists have characterized academic writing as especially conservative and resistant to historical change, contrasted with popular written registers that are supposedly more receptive to change. This stereotype persists because linguists have assumed that all historical change originates in speech, and thus they have considered only linguistic innovations that originated in speech. With that restricted focus, the stereotype is accurate: the findings reported in [Chapter 4](#) (see [Table 4.2](#)) show that colloquial innovations like semi-modals and phrasal verbs are adopted quickly in popular written registers like fiction, while they have generally not been adopted in academic prose. Thus, academic writing is indeed resistant to change, *if* we consider only grammatical features that originate in speech.

However, when we entertain the possibility that grammatical change can originate in writing, we start to notice completely different kinds of linguistic features, and we discover that academic writing has been anything but conservative and resistant to change. In fact, academic writing has been the leader in these other kinds of historical developments, employing grammatical patterns of use that are only gradually becoming adopted in popular written registers. When the full suite of phrasal complexity features are considered, we see a dramatic historical shift in grammatical discourse style that is unprecedented in any other register: moving away from the heavy reliance on verbs and dependent clauses in the eighteenth century, to the predominance of noun phrases and phrasal modifiers (coupled with the relatively rare use of verbs and dependent clauses) in the late twentieth century.

As noted in [Section 7.1](#), many linguists operate with the stereotype that grammatical complexity is equivalent to elaborated grammar, which can be operationalized as the dense use of dependent clauses. Because academic writing is complex in content and takes a long time to learn, it has simply been assumed that it must be structurally elaborated – and thus, must employ many dependent clauses. We have shown that this stereotype is also fundamentally incorrect: academic writing is indeed complex, but it is structurally compressed rather than elaborated, and the predominant grammatical complexities of academic writing are phrasal rather than clausal. Recognition of these distinctive grammatical characteristics allows us to trace the dramatic historical changes from the elaborated clausal style of eighteenth century academic research articles to the compressed phrasal style of modern research articles.

An associated stereotype is that academic writing is maximally explicit in the expression of meaning. The primary communicative purpose of academic writing is to convey information, but there is normally no direct contact between author and reader, and thus no chance to clarify meanings that are unclear. Given that situation, it seems obvious that academic writers should employ grammatical structures that are maximally explicit in meaning. But they do not. In fact, the compressed phrasal devices employed in academic writing could be described as maximally *inexplicit* in meaning. These devices omit verbs and grammatical signals that would help readers identify the meaning relations among elements. The resulting compressed structures can be highly

efficient for expert readers, but they are certainly not explicit in the expression of meaning.¹

All of these peculiar linguistic patterns are especially characteristic of specialist science research writing. Academic writing has often been treated as a single general register, making it easy to disregard the systematic patterns of linguistic variation among academic sub-registers. However, consideration of academic sub-registers is crucially important for recognizing the distinctive linguistic patterns of linguistic variation documented in the preceding chapters. At one extreme, humanities research writing *is* relatively conservative and resistant to change: to a large extent, modern humanities books and articles employ the same grammatical discourse styles as eighteenth century academic prose, relying heavily on elaborated structures with dependent clauses. In contrast, specialist science research articles have changed radically in their grammatical discourse style, and the innovative phrasal complexity features and the dramatic patterns of historical change documented in preceding chapters are especially characteristic of that academic sub-register.

The discussion in [Sections 7.1–7.3](#) have primarily focused on the theoretical implications of the findings presented in this book, challenging commonly-held perspectives on the nature of academic writing, the study of linguistic change, and the nature of grammatical complexity in English.

¹ The positive value attached to a compressed and inexplicit discourse style in academic research writing has in some cases affected the preferred rhetorical style of research articles, even resulting in overt prescriptions in the style sheets of some academic journals. One notable example of this type comes from the style sheet for the APA (*American Psychological Association*), which has been widely adopted by many research journals in the social sciences (including applied linguistics). In most applied linguistics journals, this style sheet prohibits the use of numbered sections in research articles. Although numbering conventions have no semantic content, they do provide an overt indication of the hierarchical logical relations among the sections of a text. However, in research articles conforming to APA requirements, the logical relations among sections are shown implicitly, through typographical devices such as italics, capitalization, and alignment. Such implicit marking of the rhetorical relations in an article works reasonably well for a prototypical (and stereotypical) experimental study, which has only four main sections: Introduction–Methods–Results–Discussion/Conclusion. However, it becomes very difficult even for expert readers to decipher the intended logical relationships among sections when there are sub-sections and multiple levels of headings.

In general, we have adopted a value-neutral approach in our descriptions of the compressed grammatical styles found in present-day academic writing. We have shown how the preferred grammatical features of modern academic prose are structurally compressed and inexplicit in meaning, but we have not taken an overt stance on the desirability of such features. However, we consider the APA prescription against numbered sections as fundamentally different: it prohibits the explicit marking of logical organization in a text, with essentially no savings in structural compression. We find it disconcerting that applied linguists, who should have the most expertise in the discourse factors that facilitate reading comprehension, have uncritically adopted a style sheet that prohibits the explicit marking of meaning relationships among rhetorical sections in a research article. Although this trend conforms to the general historical drift towards less explicit styles in academic prose, it further indicates that this discourse style has become aesthetically preferred as an end in itself, apart from any functional benefits that might be associated with the reliance on compressed structures.

However, these linguistically- and research-oriented discussions naturally prompt a consideration of the resulting implications for more applied areas of linguistic research. We turn to a brief discussion of these applied concerns in [Section 7.4](#), demonstrating a range of connections between the empirical patterns documented throughout the book and the interests of applied linguistics working in language teaching, learning, and assessment.

7.4 Implications for applied linguists: From language development to language teaching and assessment

The linguistic findings presented here also challenge the stereotypes and practices of applied linguists. For example, studies of language development (especially writing development) have been based on the assumptions that grammatical complexity is a unified construct, that it can be measured through a single omnibus measure (usually based on ‘t-units’), and that the increased use of dependent clauses represents increased proficiency. Practices for language teaching and assessment are based on these same assumptions. In contrast, the compressed discourse styles actually found in advanced academic writing are often disregarded in the teaching and testing of academic reading and writing.

Taking a register variation perspective, coupled with detailed consideration of phrasal complexity structures, leads to fundamentally different expectations about language development. We have discussed the applied implications of our research in several previous publications (e.g., Biber and Gray 2010; Biber, Gray, and Poonpon 2011, 2013; Biber, Gray, and Staples 2014). For experts in an academic sub-discipline, the compact, inexplicit discourse style of research articles is functional and efficient. By compressing information into few words, specialists can quickly read through a text and obtain the information they need. The lack of explicitness rarely causes problems, because specialists have extensive background knowledge in the subject area. However, the situation is the opposite for novice students. Students lack the specialist knowledge that would allow them to readily infer the expected meaning of compact, inexplicit grammatical constructions. Thus, there are clear implications of these findings for the teaching of academic reading: whether we like these discourse styles or not, it is a certainty that students will encounter professional written texts with these characteristics, and that students’ academic success will depend on their ability to extract the intended meaning from these texts.

Traditional grammar practice for advanced academic purposes tends to focus on elaborated structures, usually dependent clauses like relative clauses, adverbial clauses, and complement clauses. Phrasal modification – especially noun phrase structures with multiple levels of phrasal embedding – is given much less attention. This practice might in part reflect a perception that phrasal modification is somehow easier than clausal embedding. However, as documented in the preceding chapters, the opposite is in fact the case: meaning relations are much clearer and more explicitly expressed through clausal embedding than phrasal embedding. The applied implication is that English for academic purposes (EAP) reading courses should provide extensive practice in decoding texts with extensive phrasal modification (see, e.g., Musgrave and Parkinson 2014 on teaching nouns as nominal pre-modifiers).

The applied implications for academic writing instruction are more complicated. When teaching advanced academic writing, both ‘clarity’ and ‘economy’ are important. There is a fundamental tension between these competing goals. Students need to learn how to write prose that is clear, where the reader has no doubts about the intended meaning. But at the same time, compact prose is valued in advanced academic writing (especially science writing), and a high level of implicitness is tolerated by expert readers. It is not clear how these competing demands are best handled in advanced EAP writing courses, but given that nearly all academic research writing has these characteristics, this is clearly an area that requires additional research.

In Biber, Gray, and Poonpon (2011: 30–31), we hypothesized a series of developmental stages for students of advanced academic reading and writing, mirroring the progression from the complexity features that are common in conversation to the complexity features that are common in academic writing. Table 7.1 summarizes the grammatical features associated with each stage. As in the present book, the hypothesized stages are based on consideration of both grammatical form and syntactic function. Thus, the stages generally progress from finite dependent clauses functioning as constituents in other clauses, through intermediate stages of non-finite dependent clauses and phrases functioning as constituents in other clauses, and finally to stages requiring dense use of phrasal (non-clausal) dependent structures that function as constituents in noun phrases.

In addition, we considered lexico-grammatical factors in positing the developmental stages, because dependent clause structures in conversation tend to occur with just a few controlling words, while structures in academic writing occur with a much wider range of controlling words. For example, 75% of all *that*-complement clauses in conversation occur

Table 7.1 *Hypothesized developmental stages for complexity features*

Stage	Grammatical structure(s)	Example(s)
Stage 1	Finite complement clauses (<i>that</i> and <i>WH</i>) controlled by extremely common verbs (e.g., <i>think</i> , <i>know</i> , <i>say</i>)	1a <i>We never quite know <u>what to make of him</u></i> (conv) 1b <i>I think I found her on the eighth floor</i> (conv)
Stage 2	Finite complement clauses controlled by a wider set of verbs Finite adverbial clauses Non-finite complement clauses, controlled by common verbs (especially <i>want</i>) Phrasal embedding in the clause: simple adverbs as adverbials Simple phrasal embedding in the noun phrase: common attributive adjectives	2a <i>I'd forgotten <u>that he had just testified on that one</u></i> (conv) 2b <i>There's some here <u>if you want it</u></i> (conv) 2c <i>I'm assuming I gained weight <u>because things are a little tighter than they used to be</u></i> (conv) 2d <i>I don't want <u>to fight with them about it</u></i> (conv) 2e <i>I bate <u>watching</u> the people interact</i> (conv) 2f <i>We came <u>here and then</u> parked <u>inside</u></i> (conv) 2g <i>It certainly has a <u>nice</u> flavor</i> (conv) 2h <i>It's one of those <u>big brown</u> trays</i> (conv)
Stage 3	Phrasal embedding in the clause: prepositional phrases as adverbials Finite complement clauses controlled by adjectives Non-finite complement clauses controlled by a wider set of verbs <i>that</i> relative clauses, especially with animate head nouns Simple phrasal embedding in the noun phrase: nouns as pre-modifiers Possessive nouns as pre-modifiers of phrases as post-modifiers Simple PPs as NP post-modifiers, especially with prepositions other than <i>of</i> when they have concrete/locative meanings	3a <i>He seems to have been hit <u>on the head</u></i> (fict) 3b <i>I was sure <u>that I could smooth over our little misunderstanding</u></i> (fict) 3c <i>The snow began <u>to fall</u> again</i> (fict) 3d <i>the guy <u>that made that call</u></i> (fict) 3e <i>a really obscure <u>cable</u> channel</i> (fict) 3f <i><u>Tobie's</u> voice</i> (fict) 3g <i>the editor <u>of the food section</u></i> (fict) 3h <i>a house <u>in the suburbs</u></i> (fict)
Stage 4	Non-finite complement clauses controlled by adjectives Extraposed complement clauses	4a <i>These will not be easy <u>to obtain</u></i> (acad) 4b <i>It is clear <u>that much remains to be learned</u></i> . . . (acad) 4c <i>In that case it is useful <u>to phrase sustainability in terms of</u></i> . . . (acad)

Table 7.1 (*cont.*)

Stage	Grammatical structure(s)	Example(s)
Stage 5	Non-finite relative clauses	4d <i>the method used here</i> (acad) 4e <i>studies employing electrophysiological measures</i> (acad)
	Multiple pre-modifiers in the noun phrase: attributive adjectives and nouns as pre-modifiers	4f <i>the prevalence of airway obstruction and self-reported disease status</i> (acad) 4g <i>Positive propagule size effects have been demonstrated for both plant and animal systems</i> (acad)
	PPs as noun post-modifiers, especially with prepositions other than <i>of</i> when they have abstract meanings	4h <i>half of the subjects in each age/instructional condition</i> (acad) 4i <i>the specific growth rate at small population sizes</i>
	Preposition + non-finite complement clause	5a <i>the idea of using a Monte Carlo approach</i> (acad)
	Complement clauses controlled by nouns	5b <i>the hypothesis that female body weight was more variable</i> (acad) 5c <i>a plan to upgrade the airport</i> (acad)
	Appositive noun phrases	5d <i>The CTBS (the fourth edition of the test) was administered in 1997–1998</i> (acad)
	Extensive phrasal embedding in the NP: multiple pre-nominal and post-nominal phrasal modifiers	5e <i>the [presence of layered [[structures] at the [[[borderline]]] of cell territories]]]</i> (acad)

with only four verbs: *think, say, know, and guess* (see Biber et al. 1999: 668). *To*-clauses are generally not frequent in conversation, but the combination *want* + *to*-clause is extremely common (see Biber et al. 1999: 711). We thus hypothesize that these relatively fixed lexico-grammatical combinations are acquired at an earlier stage than full control of the target syntactic structure.

Recent studies of writing development provide support for these hypothesized progressions. For example, Taguchi et al. (2013) use many of these same features to investigate L2-English student writing, finding that lower-rated essays used more finite/non-finite dependent clauses while higher-rated essays used more attributive adjectives and post-noun-modifying prepositional phrases. Parkinson and Musgrave (2014) utilize a subset

of the developmental stages from Biber et al. (2011) to compare L2 writing by EAP students and matriculated MA students. The results provide strong support for certain aspects of the hypothesized developmental progression at advanced stages of writing development. For example, EAP students used many more attributive adjectives than MA students, while MA students used twice as many pre-modifying nouns as EAP students (see Parkinson and Musgrave 2014: figure 1, p. 56). Similarly, post-nominal prepositional phrases were much more common in MA writing than in EAP writing.

Biber, Gray, and Staples (2014) investigated a large set of clausal and phrasal complexity features within the context of a standardized language exam (TOEFL-iBT). That study failed to find systematic differences among score levels for the use of individual complexity features. However, the first parameter of variation identified in a multi-dimensional analysis provided strong support for the basic opposition between phrasal versus clausal styles of complexity: the 'positive' features on that dimension included nouns, attributive adjectives, pre-modifying nouns, and prepositional phrases; the opposing 'negative' features on that dimension included verbs, *that*-complement clauses, and finite adverbial clauses (see Biber, Gray, and Staples 2014: table 5). The differences among score levels (and task types) were highly systematic with respect to this linguistic dimension, with higher proficiency test-takers employing more nominal/phrasal discourse styles than lower proficiency test-takers.

The results of these studies suggest that the entire suite of complexity features work in concert. Thus, while it might be difficult to isolate progression in the use of an individual feature, it is more likely that there will be important differences in the use of the full set of co-occurring complexity features. In addition, comparison of the Biber et al. and Parkinson and Musgrave studies suggests that these developments are best observed at very advanced stages of writing development. Thus, there appear to be only minor differences among proficiency levels in beginning EAP courses and undergraduate composition courses, because few students employ phrasal discourse styles at that level. At more advanced levels, though, we begin to observe the predicted developmental progressions in the use of phrasal complexity features.

Obviously, much more research is required to investigate these possibilities. The present book has provided the linguistic foundation for that applied research: challenging the profession to rethink notions of grammatical complexity and incorrect stereotypes about professional academic writing. Future research is required to track the actual paths that students

follow as they move from competence in the use of clausal complexity features found in conversation to competence in the use of phrasal complexity features found in advanced academic writing. Building on that applied research, we expect that future practice in language teaching and testing can better account for the actual patterns of language use found in advanced academic writing.

7.5 Conclusion

Our goal in the present book has been to show how the corpus-based analysis of historical linguistic change in academic writing challenges numerous widely-held beliefs and stereotypes. Rather than being restricted to spoken discourse, we have shown how major historical innovations and developments can occur in writing. Rather than being conservative and resistant to change, we have shown how specialist academic research writing has been the locus of some of the most dynamic grammatical changes in the last two centuries. For the most part, those changes involve the development of phrasal complexity features. Thus, rather than focusing on dependent clauses as indicators of grammatical complexity, we have argued that phrasal modifiers are equally important – although fundamentally different – reflections of complexity. As a result, rather than being structurally elaborated, we have shown that the hallmark grammatical style of academic research writing is characterized by structural compression. And finally, rather than being maximally explicit in the expression of meaning, we have shown how academic research writing is often amazingly inexplicit, due to the loss of meaning associated with the use of compressed rather than elaborated structures.

The corpus-based research findings in this book provide strong support for these alternative perspectives. It is our hope that the volume will serve as a starting point: a foundation for future research by scholars coming from a range of perspectives, including theoretical and descriptive linguists, discourse analysts and applied linguists, sociolinguists and historical linguists.

References

- Aijmer, Karin 2009. 'Seem and evidentiality', *Functions of Language* 16(1): 63–88.
- Aktas, Rahime Nur and Cortes, Viviana 2008. 'Shell nouns as cohesive devices in published and ESL student writing', *Journal of English for Academic Purposes* 7: 3–14.
- Altenberg, Bengt and Granger, Sylviane 2001. 'The grammatical and lexical patterning of MAKE in native and non-native student writing', *Applied Linguistics* 22(2): 173–195.
- Archer, Arlene 2008. "The place is suffering": Enabling dialogue between students' discourses and academic literacy conventions in engineering', *English for Specific Purposes* 27: 255–266.
- Aronoff, Mark 1985. 'Orthography and linguistic theory: The syntactic basis of Masoretic Hebrew Punctuation', *Language* 61(1): 28–72.
- Atkinson, Dwight 1999. *Scientific Discourse in Sociohistorical Context: The Philosophical Transactions of the Royal Society of London, 1675–1975*. Mahwah, NJ: Lawrence Erlbaum.
- Banks, David 2005. 'On the historical origins of nominalized process in scientific text', *English for Specific Purposes* 24: 347–357.
2008. *The Development of Scientific Writing: Linguistic Features and Historical Context*. London: Equinox.
- Baratta, Alexander 2009. 'Revealing stance through passive voice', *Journal of Pragmatics* 41: 1406–1421.
- Barber, Charles 1964. *Linguistic Change in Present-day English*. London and Edinburgh: Oliver and Boyd.
- Basturkmen, Helen 2009. 'Commenting on results in published research articles and masters dissertations in Language Teaching', *Journal of English for Academic Purposes* 8: 241–251.
- Bazerman, Charles 1988. *Shaping Written Knowledge*. Madison, WI: University of Wisconsin Press.
- Berkenkotter, Carol and Huckin, Thomas 1995. *Genre Knowledge in Disciplinary Communication: Cognition/Culture/Power*. Hillsdale, NJ: Lawrence Erlbaum.
- Bhatia, Vijay 1997. 'Genre-mixing in academic introductions', *English for Specific Purposes* 16(3): 181–195.
- Biber, Douglas 1985. 'Investigating macroscopic textual variation through multifeature/multi-dimensional analyses', *Linguistics* 23: 337–360.

1986. 'Spoken and written textual dimensions in English: Resolving the contradictory findings', *Language* 62: 384–414.
1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
1995. *Dimensions of Register Variation: A Cross-linguistic Comparison*. Cambridge: Cambridge University Press.
2001. 'Dimensions of variation among 18th century registers', in H.-J. Diller and M. Gortlach (eds.), *Towards a History of English as a History of Genres*, pp. 89–110. Heidelberg: C. Winter. [Reprinted in Conrad & Biber (eds.) 2001, pp. 200–214]
2003. 'Compressed noun phrase structures in newspaper discourse: The competing demands of popularization vs. economy', in J. Aitchison and D. Lewis (eds.), *New Media Discourse*, pp. 169–181. London: Routledge.
2004. 'Historical patterns for the grammatical marking of stance: A crossregister comparison', *Journal of Historical Pragmatics* 5: 107–135.
- 2006a. *University Language: A Corpus-based Study of Spoken and Written Registers*. Amsterdam: John Benjamins.
- 2006b. 'Stance in spoken and written university registers', *Journal of English for Academic Purposes* 5: 97–116.
2009. 'Are there linguistic consequences of literacy? Comparing the potentials of language use in speech and writing', in D. Olson and N. Torrance (eds.), *Cambridge Handbook of Literacy*, pp. 75–91. Berlin: Wissenschaftlicher Verlag Trier.
2012. 'Register as a predictor of linguistic variation', *Corpus Linguistics and Linguistic Theory* 8: 9–37.
2014. Using multi-dimensional analysis to explore cross-linguistic universals of register variation. *Languages in Contrast* 14(1): 7–34.
- Biber, Douglas and Clark, Victoria 2002. 'Historical shifts in modification patterns with complex noun phrase structures: How long can you go without a verb?', in T. Fanego, M. José López-Couso and J. Pérez-Guerra (eds.), *English Historical Syntax and Morphology*, pp. 43–66. Amsterdam: John Benjamins.
- Biber, Douglas and Conrad, Susan 2009. *Register, Genre, and Style*. Cambridge: Cambridge University Press.
- Biber, Douglas, Conrad, Susan and Cortes, Viviana 2004. 'If you look at...: Lexical bundles in university teaching and textbooks', *Applied Linguistics* 25: 371–405.
- Biber, Douglas, Conrad, Susan and Reppen, Randi 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Biber, Douglas, Egbert, Jesse, Gray, Bethany, Oppliger, Rahel and Szmrecsanyi, Benedikt 2016. 'Variation versus text-linguistic approaches to grammatical change in English: Nominal modifiers of head nouns', in M. Kytö and P. Paivi (eds.), *Handbook of English Historical Linguistics*. Cambridge: Cambridge University Press.

- Biber, Douglas and Finegan, Edward 1988. 'Adverbial stance types in English', *Discourse Processes* 11: 1–34.
- 1989a. 'Drift and the evolution of English style: A history of three genres', *Language* 65: 487–517.
- 1989b. 'Styles of stance in English: Lexical and grammatical marking of evidentiality and affect', *Text* 9: 93–124.
1997. 'Diachronic relations among speech-based and written registers in English', in T. Nevalainen and L. Kahlas-Tarkka (eds.), *To Explain the Present: Studies in Changing English in Honor of Matti Rissanen*, pp. 253–276. Helsinki: Société Néophilologique. [reprinted in Conrad and Biber 2001, pp. 66–83]
- Biber, Douglas, Finegan, Edward and Atkinson, Dwight 1994. 'ARCHER and its challenges: Compiling and exploring a Representative Corpus of Historical English Registers', in U. Fries, G. Tottie and P. Schneider (eds.), *Creating and Using English Language Corpora*, pp. 1–14. Amsterdam: Rodopi.
- Biber, Douglas and Gray, Bethany 2010. 'Challenging stereotypes about academic writing: Complexity, elaboration, explicitness', *Journal of English for Academic Purposes* 9: 2–20.
2012. 'The competing demands of popularization vs. economy: Written language in the age of mass literacy', in T. Nevalainen & E. Traugott (eds.), *The Oxford Handbook of the History of English*, pp. 314–328. Oxford: Oxford University Press.
- 2013a. Discourse Characteristics of Writing and Speaking Task Types on the TOEFL iBT Test: A Lexico-grammatical Analysis. *TOEFLiBT Research Report (TOEFL iBT-19)*. Princeton, NJ: Educational Testing Service.
- 2013b. 'Being specific about historical change: The influence of sub-register', *Journal of English Linguistics* 41(2): 104–134.
- Biber, Douglas, Gray, Bethany and Poonpon, Kornwipa 2011. 'Should we use characteristics of conversation to measure grammatical complexity in L2 writing development?', *TESOL Quarterly* 45(1): 5–35.
2013. 'Pay attention to the phrasal structures: Going beyond T-units—A response to WeiWei Yang', *TESOL Quarterly* 47(1): 192–201.
- Biber, Douglas, Gray, Bethany and Staples, Shelley. 2014. Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*. (Currently available through Advanced Access: <http://applied.oxfordjournals.org/content/early/2014/10/21/applin.amuo59.short?rss=1>).
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan and Finegan, Edward 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Biber, Douglas and Jones, James 2009. 'Quantitative methods in corpus linguistics', in A. Lüdeling & M. Kytö (eds.), *Corpus Linguistics: An International Handbook*, pp. 1286–1304. Berlin: Walter de Gruyter.
- Biber, Douglas and Kurjian, Jerry 2007. 'Towards a taxonomy of web registers and text types: A multi-dimensional analysis', in M. Hundt, N. Nesselhauf,

- and C. Biewer (eds.), *Corpus Linguistics and the Web*, pp. 109–132. Amsterdam: Rodopi.
- Bloomfield, Leonard 1933. *Language*. New York: Holt, Reinhart & Winston.
- Brett, Paul. 1994. 'A genre analysis of the results section of sociology articles', *English for Specific Purposes* 13(1): 47–59.
- Brown, Annie, Iwashita, Noriko and McNamara, Tim 2005. 'An examination of rater orientations and test-taker performance on English-for-Academic-Purposes speaking tasks', *TOEFL Monograph Series MS-29*. Princeton, NJ: Educational Testing Service.
- Brown, Gillian and Yule, George 1983. *Discourse Analysis*. Cambridge: Cambridge University Press.
- Bruce, Ian 2008. 'Cognitive genre structures in Methods sections of research articles: A corpus study', *Journal of English for Academic Purposes* 7: 38–54.
2009. 'Results sections in sociology and organic chemistry articles: A genre analysis', *English for Specific Purposes* 28: 105–124.
- Bunton, David 2005. 'The structure of PhD conclusion chapters', *Journal of English for Academic Purposes* 4: 207–224.
- Bybee, Joan and Hopper, Paul (eds.) 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.
- Carter, Ronald and McCarthy, Michael 2006. *Cambridge Grammar of English*. Cambridge: Cambridge University Press.
- Chafe, Wallace 1982. 'Integration and involvement in speaking, writing, and oral literature', in D. Tannen (ed.), *Spoken and Written Language: Exploring Orality and Literacy*, pp. 35–54. Norwood, NJ: Ablex.
- Chambers, J.K. 2002. 'Patterns of variation including change', in J.K. Chambers, P. Trudgill, and N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*, pp. 349–373. Oxford: Wiley-Blackwell.
- Charles, Maggie 2003. 'This mystery...': a corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines', *Journal of English for Academic Purposes* 2: 313–326.
- 2006a. 'Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines', *English for Specific Purposes* 25: 310–331.
- 2006b. 'The construction of stance in reporting clauses: A cross-disciplinary study of theses', *Applied Linguistics* 27: 492–518.
2007. 'Argument or evidence? Disciplinary variation in the use of the Noun that pattern in stance construction', *English for Specific Purposes* 26: 203–218.
- Conrad, Susan 1996. 'Investigating academic texts with corpus-based techniques: An example from biology', *Linguistics and Education* 8: 299–326.
- Conrad, Susan and Biber, Douglas 2001. *Variation in English: Multi-dimensional Studies*. London: Longman.
- Cortes, Viviana 2004. 'Lexical bundles in published and student disciplinary writing: Examples from history and biology', *English for Specific Purposes* 23(4): 397–423.
- Coxhead, Averil 2000. 'A new academic word list', *TESOL Quarterly* 34(2): 213–238.

- Crespo García, Begoña and Moskowich-Spiegel Fandiño, Isabel 2010. 'CETA in the context of the Coruña Corpus', *Literary and Linguistic Computing* 25(2): 153–164.
- Croft, William 2000. *Explaining Language Change: An Evolutionary Approach*. London: Longman.
- Crompton, Peter 1997. 'Hedging in academic writing: Some theoretical problems', *English for Specific Purposes* 16(4): 271–287.
- Culpeper, Jonathan and Kytö, Merja 2010. *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: Cambridge University Press.
- Davies, Mark 2010. 'The Corpus of Contemporary American English as the first reliable monitor corpus of English', *Literary and Linguistic Computing* 25(4): 447–464.
2012. 'Expanding horizons in historical linguistics with the 400-million word Corpus of Historical American English', *Corpora* 7(2): 121–157.
- Denison, David 1998. 'Syntax', in S. Romaine (ed.), *The Cambridge History of the English Language*, vol. IV, 1776–1997, pp. 92–329. Cambridge: Cambridge University Press.
- DeVito, Joseph 1966. 'Psychogrammatical factors in oral and written discourse by skilled communicators', *Speech Monographs* 33: 73–76.
- Diani, Giuliana 2008. 'Emphasizers in spoken and written academic discourse: The case of really', *International Journal of Corpus Linguistics* 13(3): 296–321.
- Dressen, Dacia 2003. 'Geologists' implicit persuasive strategies and the construction of evaluative evidence', *Journal of English for Academic Purposes* 2: 273–290.
- Ellis, Nick, Simpson-Vlach, Rita and Maynard, Carson 2008. 'Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL', *TESOL Quarterly* 42(3): 375–396.
- Ellis, Rod and Yuan, Fangyuan 2004. 'The effects of planning on fluency, complexity, and accuracy in second language narrative writing', *Studies in Second Language Acquisition* 26: 59–84.
- Fairclough, Norman 1992. *Discourse and Social Change*. Cambridge: Polity.
- Fang, Zhihui, Schleppergrell, Mary and Cox, Beverly 2006. 'Understanding the language demands of schooling: Nouns in academic registers', *Journal of Literacy Research* 38(3): 247–273.
- Feist, Jim 2012. *Premodifiers in English: Their Structure and Significance*. Cambridge: Cambridge University Press.
- Fennell, Barbara 2001. *A History of English*. Malden, MA: Blackwell.
- Ferguson, Gibson 2001. 'If you pop over there: A corpus-based study of conditionals in medical discourse', *English for Specific Purposes* 20(1): 61–82.
- Fillmore, Charles. 1981. 'Pragmatics and the description of discourse', in P. Cole (ed.), *Radical Pragmatics*, pp. 143–166. New York: Academic Press.
- Finegan, Edward 2008. *Language: Its Structure and Use* (5th ed). Boston, MA: Thompson Wadsworth.
- Flowerdew, John (ed.) 2002. *Academic Discourse*. New York: Longman.

2006. 'Use of signaling nouns in a learner corpus,' *International Journal of Corpus Linguistics* 11(3): 345–362.
- Flowerdew, Lynn 2002. 'Corpus-based analyses in EAP', in J. Flowerdew (ed.), *Academic Discourse*, pp. 95–115.
- Fortanet, Inmaculada 2008. 'Evaluative language in peer review referee reports', *Journal of English for Academic Purposes* 7: 27–37.
- Fox, Barbara 2007. 'Principles shaping grammatical practices', *Discourse Studies* 9: 299–318.
- Freddi, Maria 2005. 'Arguing linguistics: Corpus investigation of one functional variety of academic discourse', *Journal of English for Academic Purposes* 4: 5–26.
- Gardner, Dee and Davies, Mark 2014. 'A new academic vocabular list', *Applied Linguistics* 35(3): 305–327.
- Gardner, Sheila and Nesi, Hilary 2012. 'A classification of genre families in university student writing', *Applied Linguistics* 34(1): 1–29.
- Gilbert, G. Nigel and Mulkay, Michael 1984. *Opening Pandora's Box: A Sociological Analysis of Scientific Discourse*. Cambridge: Cambridge University Press.
- Gledhill, Chris 2000. 'The discourse function of collocation in research article introductions', *English for Specific Purposes* 19(2): 115–135.
- Grabe, William and Kaplan, Robert 1996. *Theory and Practice of Writing: An Applied Linguistic Perspective*. London: Longman.
1997. 'On the writing of science and the science of writing: Hedging in science text and elsewhere', in R. Markkanen and H. Schroder (eds.), *Hedging and Discourse: Approaches to the Analysis of a Pragmatic Phenomenon in Academic Texts*, pp. 151–167. Berlin: Walter de Gruyter.
- Grant, Leslie and Ginther, April 2000. 'Using computer-tagged linguistic features to describe L2 writing differences', *Journal of Second Language Writing* 9(2): 123–145.
- Gray, Bethany 2010. 'On the use of demonstrative pronouns and determiners as cohesive devices: A focus on sentence-initial *this/these* in academic prose', *Journal of English for Academic Purposes* 9: 167–183.
2013. 'More than discipline: Uncovering multi-dimensional patterns of variation in academic research articles', *Corpora* 8(2): 153–181.
- 2015a. 'On the complexity of academic writing: Disciplinary variation and structural complexity', in V. Cortes and E. Csomay (eds.), *Corpus-based Research in Applied Linguistics. In Honor of Douglas Biber*, pp. 49–77. Amsterdam: John Benjamins.
- 2015b. *Linguistic Variation in Research Articles: When Discipline Tells Only Part of the Story*. Amsterdam: John Benjamins.
- Gray, Bethany and Cortes, Viviana 2010. 'Perception vs. evidence: An analysis of this and these in academic prose', *English for Specific Purposes* 30(1): 31–43.
- Green, Christopher, Christopher, Elsie and Mei, Jaquelin Lam Kam 2000. 'The incidence and effects on coherence of marked themes in interlanguage texts: A corpus-based enquiry', *English for Specific Purposes* 19: 99–113.

- Grieve, Jack, Biber, Douglas, Friginal, Eric and Nekrasova, Tatiana 2011. 'Variation among blogs: A multi-dimensional analysis', in A. Mehler, S. Sharoff and M. Santini (eds.), *Genres on the Web: Computational Models and Empirical Studies*, pp. 303–322. Springer.
- Groom, Nicholas 2005. 'Pattern and meaning across genres and disciplines: An exploratory study', *Journal of English for Academic Purposes* 4: 257–277.
- Gross, Alan, Harmon, Joseph and Reidy, Michael 2002. *Communicating Science: The Scientific Article from the 17th Century to the Present*. Oxford: Oxford University Press.
- Hall, Robert 1964. *Introductory Linguistics*. Philadelphia: Chilton Books.
- Halliday, M.A.K. 1979. 'Differences between spoken and written language: Some implications for language teaching', in G. Page et al. (eds.), *Communication through Reading: Proceedings of the 4th Australian Reading Conference*, pp. 37–52. Adelaide: Australian Reading Association.
- Halliday, M.A.K. and Martin, J.R. 1993/1996. *Writing Science: Literacy and Discursive Power*. London: Falmer Press.
- Halliday, M.A.K. 1988. 'On the language of physical science', in M. Ghadessy (ed.), *Registers of Written English*, pp. 162–178. London: Pinter. [reprinted in M.A.K. Halliday 2004]
2004. *The Language of Science*. London: Continuum.
- Halliday, M.A.K. and Martin, J.R. 1993. *Writing Science: Literacy and Discursive Power*. Pittsburgh: University of Pittsburgh Press.
- Hardy, Jack and Römer, Ute 2013. 'Revealing disciplinary variation in student writing: A multi-dimensional analysis of the Michigan Corpus of Upper-Level Student Papers (MICUSP)', *Corpora* 8(2): 183–207.
- Hewings, Ann, Lillis, Theresa and Vladimirov, Dimitra 2010. 'Who's citing whose writings? A corpus based study of citations as interpersonal resource in English medium national and English medium international journals', *Journal of English for Academic Purposes* 9: 83–150.
- Hewings, Martin (ed.) 2001. *Academic Writing in Context: Implications and Applications*. Birmingham: The University of Birmingham Press.
- Hewings, Martin and Hewings, Ann 2001. 'Anticipatory "it" in academic writing: An indicator of disciplinary difference and developing disciplinary knowledge', in M. Hewings (ed.), *Academic Writing in Context: Implications and Applications*, pp. 199–214.
2002. "It is interesting to note that...": A comparative study of anticipatory 'it' in student and published writing', *English for Specific Purposes* 21: 367–383.
- Hinkel, Eli 2003. 'Simplicity without elegance: Features of sentences in L1 and L2 academic texts', *TESOL Quarterly* 37(2): 275–301.
- Hinrichs, Lars and Szmrecsanyi, Benedikt 2007. 'Recent changes in the function and frequency of standard English genitive constructions: A multivariate analysis of tagged corpora', *English Language and Linguistics* 11(3): 437–474.
- Hoffmann, Sebastian 2005. *Grammaticalization and English Complex Prepositions: A Corpus-based Study*. London: Routledge.

- Holmes, Janet 1986. 'Doubt and certainty in ESL textbooks', *Applied Linguistics* 9: 21–43.
1992. *An Introduction to Sociolinguistics*. Harlow, England: Longman.
- Hopper, Paul J. and Traugott, Elizabeth C. 2003. *Grammaticalization* (2nd edition). Cambridge: Cambridge University Press.
- Huckin, Thomas and Pesante, Linda Hutz 1988. 'Existential there', *Written Communication* 5(3): 368–391.
- Huckin, Thomas, Haynes, Margot and Coady, James (eds.) 1995. *Second Language Reading and Vocabulary Learning*. Norwood, NJ: Ablex Publishing Corporation.
- Huddleston, Rodney 1984. *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.
- Huddleston, Rodney and Pullum, Geoffrey K. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hughes, Rebecca 1996. *English in Speech and Writing: Investigating Language and Literature*. London: Routledge.
- Hundt, Marianne 2004. 'Animacy, agentivity, and the spread of the progressive in Modern English', *English Language and Linguistics* 8: 47–69.
- Hundt, Marianne and Mair, Christian 1999. "Agile" and "uptight" genres: The corpus-based approach to language change in progress', *International Journal of Corpus Linguistics* 4: 221–242.
- Hunston, Susan 1993. 'Evaluation and ideology in scientific writing', in M. Ghadessy (ed.), *Register Analysis: Theory and Practice*, pp. 57–73. London: Pinter.
1994. 'Evaluation and organization in a sample of written academic discourse', in M. Coulthard (ed.), *Advances in Written Text Analysis*, pp. 191–218. London: Routledge.
- Hyland, Ken 1994. 'Hedging in academic writing and EAP textbooks', *English for Specific Purposes* 13(3): 239–256.
- 1996a. 'Talking to the academy: Forms of hedging in science research articles', *Written Communication* 13(2): 251–281.
- 1996b. 'Writing without conviction? Hedging in science research articles', *Applied Linguistics* 17(4): 433–454.
- 1998a. 'Boosting, hedging and the negotiation of academic knowledge', *Text* 18(3): 349–382.
- 1998b. *Hedging in Scientific Research Articles*. Amsterdam: John Benjamins.
1999. 'Academic attribution: Citation and the construction of disciplinary knowledge', *Applied Linguistics* 20(3): 341–367.
2001. 'Bringing in the reader: Addressee features in academic articles', *Written Communication* 18(4): 549–574.
- 2002a. 'Directives: Argument and engagement in academic writing', *Applied Linguistics* 23(2): 215–239.
- 2002b. 'What do they mean? Questions in academic writing', *Text* 22(4): 529–557.
- 2002c. 'Authority and invisibility: Authorial identity in academic writing', *Journal of Pragmatics* 34(8): 1091–1112.

2007. 'Applying a gloss: Exemplifying and reformulating in academic discourse', *Applied Linguistics* 28(2): 266–285.
2008. 'As can be seen: Lexical bundles and disciplinary variation', *English for Specific Purposes* 27: 4–21.
- Hyland, Ken and Tse, Polly 2005. 'Hooking the reader: A corpus study of evaluative that in abstracts', *English for Specific Purposes* 24: 123–139.
- Jarvis, Scott, Grant, Leslie, Bikowski, Dawn and Ferris, Dana 2003. 'Exploring multiple profiles of highly rated learner compositions', *Journal of Second Language Writing* 12: 377–403.
- Johns, Ann 1997. *Text, Role, and Context: Developing Academic Literacies*. Cambridge: Cambridge University Press.
- Kanoksilapatham, Budsaba 2005. 'Rhetorical structure of biochemistry research articles', *English for Specific Purposes* 24: 269–292.
- Kay, Paul 1977. 'Language evolution and speech style', in B. Blount and M. Sanches (eds.), *Sociocultural Dimensions of Language Change*, pp. 21–33. New York: Academic Press.
- Keen, John 2004. 'Sentence-combining and redrafting processes in the writing of secondary school students in the UK', *Linguistics and Education* 15: 81–97.
- Kemmer, Suzanne and Barlow, Michael 2000. 'Introduction: A usage-based conception of language', in M. Barlow and S. Kemmer (eds.), *Usage Based Models of Language*, pp. vii–xxviii. Chicago, IL: The University of Chicago Press.
- Koutsantoni, Dimitra 2006. 'Rhetorical strategies in engineering research articles and research theses: Advanced academic literacy and relations of power', *Journal of English for Academic Purposes* 5: 19–36.
- Kroll, Barbara 1977. 'Ways communicators encode propositions in spoken and written English: A look at subordination and coordination', in E. O. Keenan and T. Bennett (eds.), *Discourse across Time and Space (SCOPIL no. 5)*, pp. 69–108. Los Angeles: University of Southern California.
- Krug, Manfred 2000. *Emerging English Modals: A Corpus-based Study of Grammaticalization*. Berlin: Mouton de Gruyter.
- Labov, William 1994. *Principles of Linguistic Change*, vol. I, Internal Factors. Oxford: Blackwell.
2001. *Principles of Linguistic Change*, vol. II, Social Factors. Oxford: Blackwell.
2010. *Principles of Linguistic Change*, vol. III, Cognitive and Cultural Factors. Oxford: Blackwell.
- Langacker, Ronald 1987. *Foundations of Cognitive Grammar*. Stanford University Press.
- Larsen-Freeman, Diane 2006. 'The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English', *Applied Linguistics* 27: 590–619.
- Leech, Geoffrey 2003. 'Modal on the move: The English modal auxiliaries 1961–1992', in R. Facchinetti, F. Palmer and M. Krug (eds.), *Modality in Contemporary English*, pp. 223–240. Berlin: Mouton de Gruyter.

2011. 'The modals ARE declining. Reply to Neil Millar's "Modal verbs in TIME: The English modal auxiliaries 1961–1991"', *International Journal of Corpus Linguistics* 16(4): 547–564.
- Leech, Geoffrey, Hundt, Marianne, Mair, Christian and Smith, Nicholas 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.
- Lewin, Beverly 2005. 'Contentiousness in science: The discourse of critique in two sociology journals', *Text* 25: 723–744.
- Li, Li-Juan and Ge, Guang-Chun 2009. 'Genre analysis: Structural and linguistic evolution of the English-medium medical research article (1985–2004)', *English for Specific Purposes* 28: 93–104.
- Li, Yili 2000. 'Linguistic characteristics of ESL writing in task-based e-mail activities', *System* 28: 229–245.
- Lim, Jason Miin Hwa 2006. 'Method sections of management research articles: A pedagogically motivated qualitative study', *English for Specific Purposes* 25: 282–309.
- Lindquist, Hans and Mair, Christian 2004. *Corpus Approaches to Grammaticalization in English*. Amsterdam: John Benjamins.
- Loudermilk, Barbara Conner 2007. 'Occluded academic genres: An analysis of the MBA thought essay', *Journal of English for Academic Purposes* 6: 190–205.
- Mair, Christian 1995. 'Changing patterns of complementation, and concomitant grammaticalisation, of the verb help in present-day British English', in B. Aarts and C. Meyer (eds.), *The Verb in Contemporary English: Theory and Description*, pp. 258–272. Cambridge: Cambridge University Press.
2002. 'Three changing patterns of verb complementation in Late Modern English: A real-time study based on matching text corpora', *English Language and Linguistics* 6(1): 105–131.
2006. *Twentieth-century English: History, Variation and Standardization*. Cambridge: Cambridge University Press.
- Mair, Christian and Hundt, Marianne 1995. 'Why is the progressive becoming more frequent in English? A corpus-based investigation of language change in progress', *Zeitschrift für Anglistik und Amerikanistik* 43: 111–122.
- Marco, María José Luzon 2000. 'Collocational frameworks in medical research papers: A genre-based study', *English for Specific Purposes* 19: 63–86.
- Markkanen, Raija and Schroder, Hartmut (eds.) 1997. *Hedging and Discourse: Approaches to the Analysis of a Pragmatic Phenomenon in Academic Texts*. Berlin: Walter de Gruyter.
- Mauranen, Anna and Bondi, Marina 2003. 'Evaluation in academic discourse (special issue)', *Journal of English for Academic Purposes* 2: 269–271.
- McEnery, Tony, Xiao, Richard and Tono, Yukio 2006. *Corpus-based Language Studies: An Advanced Resource Book*. New York: Routledge.
- McWhorter, John 2001. *The Word on the Street: Debunking the Myth of "Pure" Standard English*. New York: Basic Books.
- Meyer, P.G. 1997. 'Hedging strategies in written academic discourse: Strengthening the argument by weakening the claim', in R. Markkanen

- and H. Schroder (eds.), *Hedging and Discourse: Approaches to the Analysis of a Pragmatic Phenomenon in Academic Texts*, pp. 21–41. Berlin: Walter de Gruyter.
- Millar, Neil 2009. 'Modal verbs in TIME: Frequency changes 1923–2006', *International Journal of Corpus Linguistics* 14(2): 191–220.
- Moore, Tim 2002. 'Knowledge and agency: A study of 'metaphenomenal discourse' in textbooks from three disciplines', *English for Specific Purposes* 21: 347–366.
- Musgrave, Jill and Parkinson, Jean 2014. 'Getting to grips with noun groups', *ELT Journal* 68(2): 145–154.
- Myers, Greg 1989. 'The pragmatics of politeness in scientific articles', *Applied Linguistics* 10(1): 1–35.
1990. 'The rhetoric of irony in academic writing', *Written Communication* 7(4): 419–455.
- Nation, I.S.P. 1990. *Teaching and Learning Vocabulary*. New York: Newbury House.
2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nelson, Nickola and Van Meter, Adelia 2007. 'Measuring written language ability in narrative samples', *Reading and Writing Quarterly* 23(3): 287–309.
- Nesi, Hilary and Gardner, Sheila 2012. *Genres across the Disciplines: Student Writing in Higher Education*. Cambridge: Cambridge University Press.
- Nesselhauf, Nadja 2007. 'The spread of the progressive and its 'future' use', *English Language and Linguistics* 11(1): 191–207.
- Nevanlinna, Saara and Pahta, Päivi 1997. 'Middle English nonrestrictive expository apposition with an explicit marker', in J. Fisiak (ed.), *Studies in Middle English Linguistics*, pp. 373–401. Berlin: Mouton de Gruyter.
- Nevalainen, Terttu 2004. 'Three perspectives on grammaticalization: Lexicogrammar, corpora and historical sociolinguistics', in H. Lindquist and C. Mair (eds.), *Corpus Approaches to Grammaticalization in English*, pp. 1–31. Amsterdam: John Benjamins.
- Nevalainen, Terttu and Raumolin-Brunberg, Helena 1996. *Sociolinguistics and Language History: Studies Based on the Corpus of Early English Correspondence*. Amsterdam: Rodopi.
- Nichols, Johanna 1984. 'Functional Theories of Grammar', *Annual Review of Anthropology* 13: 97–117.
- Norrby, Catrin and Håkansson, Gisela 2007. 'The interaction of complexity and grammatical processability: The case of Swedish as a foreign language', *International Review of Applied Linguistics* 45: 45–68.
- Ochs, Elinor, Schegloff, Emanuel and Thompson, Sandra 1996. *Interaction and Grammar*. Cambridge: Cambridge University Press.
- O'Donnell, Roy C. 1974. 'Syntactic differences between speech and writing', *American Speech* 49: 102–110.

- O'Donnell, Roy C., Griffin, William J. and Norris, Raymond C. 1967. *Syntax of Kindergarten and Elementary School Children: A Transformational Analysis*. Champaign, IL: NCTE.
- Oakey, David 2002. 'Formulaic language in English academic writing: A corpus-based study of formal and functional variation of a lexical phrase in different academic disciplines', in R. Reppen, S. Fitzmaurice, and D. Biber (eds.), *Using Corpora to Explore Linguistic Variation*, pp. 111–129. Philadelphia: John Benjamins.
- Olson, David R. 1977. 'From utterance to text: The bias of language in speech and writing', *Harvard Educational Review* 47: 257–281.
- Ozturk, Ismet 2007. 'The textual organization of research article introductions in applied linguistics: Variability within a single discipline', *English for Specific Purposes* 26: 25–38.
- Pahta, Päivi and Nevanlinna, Saara 1997. 'Re-phrasing in Early English: The use of expository apposition with an explicit marker from 1350–1710', In M. Rissanen, M. Kytö, and K. Heikkonen (eds.), *English in Transition: Corpus-based Studies in Linguistic Variation and Genre Styles*, pp. 121–183. Berlin: Mouton de Gruyter.
- Parkinson, Jean and Musgrave, Jill 2014. 'Development of noun phrase complexity in the writing of English for academic purposes students', *Journal of English for Academic Purposes* 14: 48–59.
- Purpura, James 2004. *Assessing Grammar*. Cambridge: Cambridge University Press.
- Raumolin-Brunberg, Helena 1991. *The Noun Phrase in Early Sixteenth-century English: A Study Based on Sir Thomas More's Writings*. Helsinki: Société Néophilologique.
- Reppen, Randi 2001. 'Register variation in student and adult speech and writing', in S. Conrad and D. Biber (eds.), *Variation in English: Multi-dimensional Studies*, pp. 187–199.
- Rissanen, M. 1999. 'Syntax', in R. Lass (ed.), *Cambridge History of the English Language*, Volume III: 1476–1776, pp. 187–331. Cambridge: Cambridge University Press.
- Rohdenburg, Gunter 1996. 'Cognitive complexity and increased grammatical explicitness in English', *Cognitive Linguistics* 7(2): 149–182.
- Römer, Ute and Swales, John 2010. 'The Michigan corpus of upper-level student papers (MICUSP)', *Journal of English for Academic Purposes* 9: 249.
- Rosenbach, Anette 2002. *Genitive Variation in English: Conceptual Factors in Synchronic and Diachronic Studies*. Berlin, New York: Mouton de Gruyter.
2007. 'Emerging variation: Determiner genitives and noun modifiers in English', *English Language and Linguistics* 11: 143–189.
- Rudanko, Juhani 2010. 'Explaining grammatical variation and change: A case study of complementation in American English over three decades', *Journal of English Linguistics* 38(1): 4–24.
- Salager-Meyer, Françoise 1994. 'Hedges and textual communicative function in medical English written discourse', *English for Specific Purposes* 13(2): 149–170.

- Samraj, Betty 2002. 'Introductions in research articles: Variations across disciplines,' *English for Specific Purposes* 21: 1–17.
2004. 'Discourse features of the student-produced academic research paper: Variations across disciplinary courses', *Journal of English for Academic Purposes* 3: 5–22.
2005. 'An exploration of a genre set: Research article abstracts and introductions in two disciplines', *English for Specific Purposes* 24: 141–156.
2008. 'A discourse analysis of master's theses across disciplines with a focus on introductions', *Journal of English for Academic Purposes* 7: 55–67.
- Sapir, Edward 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and Co.
- Schleppergrell, Mary 1996. 'Conjunction in spoken English and ESL writing', *Applied Linguistics* 17(3): 271–285.
- Schmitt, Norbert 2000. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Schmitt, Norbert and McCarthy, Michael (eds.) 1997. *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- Schneider, Edgar 2002. 'Investigating variation and change in written documents', in J.K. Chambers, P. Trudgill and N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change*, pp. 67–96. Malden, MA: Blackwell.
- Silver, Marc 2003. 'The stance of stance: A critical look at ways stance is expressed and modeled in academic discourse', *Journal of English for Academic Purposes* 2: 359–374.
- Simpson-Vlach, Rita and Ellis, Nick 2010. 'An academic formulas list: New methods in phraseology research', *Applied Linguistics* 31(4): 487–512.
- Smith, Nicholas 2002. 'Ever moving on? The progressive in recent British English', in P. Peters, P. Collins, and A. Smith (eds.), *New Frontiers of Corpus Research: Papers from the Twenty-first International Conference on English Language Research on Computerized Corpora, Sydney 2002*, pp. 317–330. Amsterdam: Rodopi.
- Smith, Nicholas and Rayson, Paul 2007. 'Recent change and variation in the British English use of the progressive passive', *ICAME Journal* 31: 129–159.
- Spycher, Pamela 2007. 'Academic writing of adolescent English learners: Learning to use "although"', *Journal of Second Language Writing* 16: 238–254.
- Strang, Barbara 1970. *A History of English*. London: Routledge.
- Staples, Shelley, Egbert, Jesse, Biber, Douglas, and Gray, Bethany in press. 'Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre'. *Written Communication*.
- Stoller, Fredricka and Robinson, Marin 2013. 'Chemistry journal articles: An interdisciplinary approach to move analysis with pedagogical aims', *English for Specific Purposes* 32: 45–57.
- Swales, John 1990. *Genre Analysis: English for Academic and Research Settings*. Cambridge: Cambridge University Press.
- Swales, John and Burke, Amy 2003. "It's really fascinating work": Differences in evaluative adjectives across academic registers', in P. Leistyna and

- C. Meyer (eds.), *Corpus Analysis: Language Structure and Language Use*, pp. 1–18. Amsterdam: Rodopi.
- Swales, John, Ahmad, Ummul K., Chang, Yu-Ying, Chavez, Daniel, Dressen, Dacia F. and Seymour, Ruth 1998. 'Consider this: The role of imperatives in scholarly writing', *Applied Linguistics* 19: 97–121.
- Szmrecsanyi, Benedikt, Biber, Douglas, Egbert, Jesse, and Franco, Karlien 2015. Two is company, three is a crowd: Modeling genitive variation in Late Modern English. Working paper.
- Tagliamonte, Sali 2004. 'Have to, gotta, must: Grammaticalisation, variation and specialization in English deontic modality', in H. Lindquist and C. Mair (eds.), *Corpus Approaches to Grammaticalization in English*, pp. 33–55. Amsterdam: John Benjamins.
2006. "'So cool, right?': Canadian English entering the 21st century', *The Canadian Journal of Linguistics* 51: 309–331.
- Taguchi, Naoko, Crawford, William, and Wetzel, Danielle Zawodny 2013. 'What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program', *TESOL Quarterly* 47: 420–430.
- Tarone, Elaine, Dwyer, Sharon, Gillette, Susan and Icke, Vincent 1998. 'On the use of the passive and active voice in astrophysics journal papers: With extensions to other language and other fields', *English for Specific Purposes* 17(1): 113–132.
- Traugott, Elizabeth 2003. 'From subjectification to intersubjectification', in R. Hickey (ed.), *Motives for Language Change*, pp. 124–139. Cambridge: Cambridge University Press.
- Tucker, Paul 2003. 'Evaluation in the art-historical research article', *Journal of English for Academic Purposes* 2: 291–312.
- Valle, Ellen 1999. 'A collective intelligence: The life sciences in the Royal Society as a scientific discourse community, 1665–1965'. Unpublished doctoral dissertation: University of Turku.
- van Gelderen, Elly 2006. *A History of the English Language*. Amsterdam: John Benjamins.
- Vande Kopple, William 1994. 'Some characteristics and functions of grammatical subjects in scientific discourse', *Written Communication* 11(4): 534–564.
- Varttala, Teppo 2003. 'Hedging in scientific research articles: A cross-disciplinary study', in G. Cortese and P. Riley (eds.), *Domain-specific English: Textual Practices across Communities and Classrooms*, pp. 141–174. New York: Peter Lang.
- Warchal, Krystyna 2010. 'Moulding interpersonal relations through conditional clauses: Consensus-building strategies in written academic discourse', *Journal of English for Academic Purposes* 9: 140–150.
- Webber, Pauline 1994. 'The function of questions in different medical journal genres', *English for Specific Purposes* 13(3): 257–268.
- Wells, Rulon 1960. 'Nominal and verbal style', in T.A. Sebeok (ed.), *Style in Language*, pp. 213–220. Cambridge: Cambridge University Press.

- Williams, Gelfrey 1998. 'Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles', *International Journal of Corpus Linguistics* 3: 151–171.
- Willis, Dave 2003. *Rules, Patterns and Words: Grammar and Lexis in English Language Teaching*. Cambridge: Cambridge University Press.
- Wolfe-Quintero, Kate, Inagaki, Shunji and Kim, Hae-Young 1998. *Second language development in writing: Measures of fluency, accuracy, and complexity* (Technical Report No. 17). Honolulu, HI: Second Language Teaching & Curriculum Center, University of Hawaii.
- Wolfram, Walt and Schilling-Estes, Natalie 2005. *American Dialects: Dialects and Variation* (2nd ed). Malden, MA: Blackwell Publishing.
- Wolk, Christoph, Bresnan, Joan, Rosenbach, Anette and Szmrecsanyi, Benedikt 2013. 'Dative and genitive variability in Late Modern English: Exploring cross-constructional variation and change', *Diachronica* 30(3): 382–419.
- Wright, Laura J. 2008. 'Writing science and objectification: Selecting, organizing, and decontextualizing knowledge', *Linguistics and Education* 19: 265–293.
- Xiao, Richard 2009. 'Multidimensional analysis and the study of world Englishes', *World Englishes* 28(4): 421–450.
- Yáñez-Bouza, Nuria 2011. 'ARCHER: Past and present (1990–2010)', *ICAME Journal* 35: 205–36.
- Yeung, Lorrita 2007. 'In search of commonalities: Some linguistic and rhetorical features of business reports as a genre', *English for Specific Purposes* 26: 156–179.

APPENDIX I

Descriptive statistics for nine linguistic features for three general registers across historical periods

Table A1.1 *Descriptive statistics for lexical verbs, adjectives, and adverbs*

Register	Period	Nouns		Lexical verbs		Adjectives		Adverbs	
		Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev
Fiction	1750	213.1	22.7	140.1	15.1	60.6	12.6	48.0	11.3
Fiction	1850	213.8	24.6	147.5	17.5	63.7	11.3	56.9	8.7
Fiction	1900	213.7	25.6	155.2	13.2	60.8	9.3	57.7	6.2
Fiction	1930	224.3	27.9	148.0	18.5	60.6	18.7	59.7	10.5
Fiction	1960	246.9	28.7	146.2	12.8	65.0	19.2	57.5	15.5
Fiction	1990	220.1	15.7	154.9	11.4	56.6	11.6	64.2	8.1
News	1750	280.0	29.8	102.8	10.4	54.4	12.6	33.6	10.4
News	1850	266.5	42.9	115.9	16.7	69.4	18.1	37.1	10.1
News	1900	301.6	33.1	121.4	19.3	58.3	18.5	35.2	11.1
News	1930	309.9	40.3	119.5	15.9	69.0	16.8	37.4	11.2
News	1960	324.7	28.2	117.9	14.7	82.7	15.7	48.4	11.4
News	1990	319.0	30.4	126.3	14.5	75.9	14.8	42.6	10.2
Science	1750	235.5	27.2	107.9	23.7	69.6	18.8	43.7	10.6
Science	1850	250.3	23.6	98.8	12.9	85.9	17.1	42.3	8.0
Science	1900	263.9	25.5	101.9	11.4	84.9	13.1	42.3	8.6
Science	1930	275.4	27.7	98.2	14.4	84.1	13.6	38.1	8.2
Science	1960	353.3	37.8	79.1	11.6	84.6	23.1	28.6	8.5
Science	1990	395.7	49.5	77.2	11.8	92.4	21.2	24.2	6.8

Table A1.2 *Descriptive statistics for passive voice verbs, nominalizations, and prepositional phrases*

Register	Period	Passive voice verbs		Nominalizations		Prepositional phrases	
		Mean	Std dev	Mean	Std dev	Mean	Std dev
Fiction	1750	13.9	6.0	36.3	11.8	117.5	16.3
Fiction	1850	10.2	5.0	26.5	7.2	108.3	14.0
Fiction	1900	8.9	3.1	25.0	6.5	103.8	11.7
Fiction	1930	9.6	3.6	20.7	8.3	106.2	15.7
Fiction	1960	10.3	3.4	25.1	8.7	103.1	9.9
Fiction	1990	9.9	1.7	19.9	4.0	101.9	8.5
News	1750	27.9	6.3	43.7	13.2	154.4	13.5
News	1850	21.1	8.8	56.2	19.6	137.5	15.9
News	1900	25.8	9.5	53.0	16.8	128.7	16.9
News	1930	20.4	8.2	51.2	16.9	124.7	15.4
News	1960	15.0	5.8	45.1	14.2	116.3	11.9
News	1990	14.2	6.6	52.9	15.0	114.1	13.7
Science	1750	24.3	9.6	37.5	14.4	143.7	19.3
Science	1850	27.5	8.4	49.3	11.9	151.2	14.9
Science	1900	24.0	7.2	53.1	12.6	145.0	13.2
Science	1930	27.8	10.6	57.0	15.6	146.8	14.1
Science	1960	24.8	7.0	53.7	14.3	126.1	18.7
Science	1990	21.7	7.7	59.4	15.3	117.0	19.3

Table A1.3 *Descriptive statistics for finite adverbial clauses, complement clauses, and relative clauses*

Register	Period	Finite adverbial clauses		Complement clauses (finite and non-finite)		Relative clauses (finite and non-finite)	
		Mean	Std dev	Mean	Std dev	Mean	Std dev
Fiction	1750	4.4	2.3	12.6	6.3	14.5	6.9
Fiction	1850	4.9	1.7	15.7	6.0	11.1	4.8
Fiction	1900	4.8	1.8	16.0	5.0	9.5	4.1
Fiction	1930	4.7	2.9	12.4	6.0	10.2	6.1
Fiction	1960	5.3	2.3	12.3	4.2	9.2	6.9
Fiction	1990	5.3	1.5	14.6	2.7	8.4	2.2
News	1750	1.8	1.8	8.6	4.1	19.3	6.9
News	1850	2.9	2.7	14.3	9.0	12.6	8.0
News	1900	3.0	2.7	14.4	8.9	13.2	8.0
News	1930	3.7	2.9	12.5	8.0	12.9	7.2
News	1960	4.3	2.8	11.5	6.6	11.8	6.4
News	1990	4.5	2.7	14.8	7.7	12.7	6.1
Science	1750	3.7	2.4	7.6	4.3	17.0	7.6
Science	1850	2.7	1.5	9.0	3.8	16.7	7.8
Science	1900	2.9	1.6	11.5	4.3	14.6	6.0
Science	1930	2.7	1.6	11.5	5.1	14.2	7.8
Science	1960	2.9	1.6	5.7	2.4	9.4	4.7
Science	1990	2.5	1.4	5.3	2.3	9.4	5.3

APPENDIX 2

Descriptive statistics for five linguistic features for sub-registers of academic writing across historical periods

Table A2.1 *Descriptive statistics for five grammatical features in academic sub-registers from three recent historical periods (1965, 1985, 2005) (rates of occurrence per 1,000 words)*

Register	Year	N	Nouns		Nominalizations		Relative clauses	
			Mean	Std dev	Mean	Std dev	Mean	Std dev
Specialist science	1965	81	353.1	37.6	53.5	13.9	4.4	2.3
	1985	64	372.5	49.1	56.2	13.7	4.2	2.5
	2005	52	423.0	26.4	61.0	16.8	4.5	2.2
Specialist social science	1965	58	337.8	32.8	64.6	13.4	6.0	2.6
	1985	58	358.1	31.9	67.6	24.9	5.5	1.9
	2005	57	372.8	36.7	70.5	13.5	6.1	2.3
Multi-disciplinary science	1985	40	312.4	30.7	61.8	20.4	7.4	2.5
	2005	40	340.9	32.5	72.1	13.0	6.6	2.1
Humanities (history)	1965	30	298.4	25.2	57.1	10.7	9.8	2.7
	1985	30	310.9	26.2	57.5	11.8	8.7	2.2
	2005	60	314.2	32.1	62.8	14.9	9.3	2.8

Register	Year	N	Noun + <i>of</i> -phrase		Noun + Noun	
			Mean	Std dev	Mean	Std dev
Specialist science	1965	81	35.7	8.6	50.3	15.1
	1985	64	33.1	7.0	60.1	19.4
	2005	52	30.2	6.0	76.6	15.1
Specialist social science	1965	58	34.8	8.3	46.2	17.0
	1985	58	34.4	7.2	58.4	14.0
	2005	57	29.8	5.0	66.3	18.6
Multi-disciplinary science	1985	40	41.4	8.5	43.2	15.1
	2005	40	35.5	5.2	56.7	13.5
Humanities (history)	1965	30	39.4	7.3	16.5	7.1
	1985	30	37.5	6.3	23.3	11.8
	2005	60	36.6	8.1	24.3	9.2

Index

- 20th Century Research Article Corpus, 53–4
- academese, 1
- academic research writing
- disciplines, 157
- academic vocabulary, 73
- academic writing, 248
- disciplines, 111
 - grammatical characteristics, 78, 83
 - historical context, 50
 - situational context, 67
 - vs. conversation, 88
- adverbial clauses, 13
- appositive noun phrases, 10, 66, 120, 132, 144, 151, 210–11, 232
- functions, 202
 - meaning relations, 233
- ARCHER, 54
- Atkinson, 44, 51, 54, 74, 158, 257, 259
- attributive adjectives, 8, 22, 132, 148, 184, 210
- classifiers, 184
 - descriptors, 184
- Banks, 76, 112, 131, 147, 257
- classroom teaching, 101
- colloquial registers, 20
- colloquialization, 128, 140, 208
- colons, 122, 241
- complement clauses, 13, 24
- compression, 18, 129, 207, 219
- conditional clauses, 74
- Conrad, 38, 49, 51, 59, 73, 78, 83, 91, 131, 247, 258–60, 268
- conversation, 88
- Corpus of English Texts on Astronomy (CETA), 56
- Corpus of Historical American English, 46
- corpus-based analyses, 43
- correlations, *See Pearson correlations*
- Davies, 44, 46, 73, 212, 214, 261–2
- definite articles, 155
- dependent clauses, 12, 16, 62, 92, 113, 141
- dependent phrases, 92, 141
- drift, 208
- economy, 129, 207
- English for Academic Purposes, 252
- explicitness, 18, 218
- Feist, 178, 216, 232, 261
- fictional novels, 132, *See novels*
- Flowerdew, 72, 262
- Fox, 15, 38, 262
- Gardner, 73, 262, 267
- genitives, 171, 180, 190, 209, 212
- grammatical complexity, 16, 62, 245
- grammatical metaphor, 130
- Gross, 50–1, 158, 263
- Halliday, 74, 76, 112, 130–1, 147, 221, 263
- Hinrichs, 45, 143, 171, 263
- historical change, 19, 27
- grammar, 28, 30
 - speech vs. writing, 32, 246
- historical context of academic research, 50
- Hopper, 31, 35, 260, 264
- Huddleston, 16, 88, 170, 264
- Hughes, 14, 88, 264
- Hundt, 20–1, 29, 33–4, 128, 166, 209, 247, 259, 264, 266, 268
- Hunston, 75, 264
- Hyland, 15, 73–5, 264
- Krug, 31, 43–4, 46, 265

- Labov, 28, 265
 Leech, 21, 28–30, 33, 43–4, 46, 48–9, 128–9, 131, 138, 171, 208, 259, 265
 Lexico-grammar, 28–9, 44, 65, 75, 252
 Lindquist, 31, 266–7, 270
 linking adverbials, 121, 240
 Longman Grammar of Spoken and Written English, 57–8, 76

 Mair, 15, 20–1, 29–31, 33–4, 43, 128, 166, 208, 247, 264, 266–7, 270
 modal verbs, 44, 137
 ‘Multi-dimensional’ (MD) analysis, 83–7

 Nesi, 73, 262, 267
 Nesselhauf, 29, 259, 267
 Nevalainen, 31, 37, 259, 267
 newspaper prose, 101, 105, 132
 Nichols, 247, 267
 nominalizations, 23, 76, 112, 125, 130, 135, 147, 161, 170, 176, 221
 noun complement clauses, 13
 noun modifiers, 94, 169
 noun phrase modifiers, 116
 noun premodifiers, 94
 noun-participle compounds, 120, 148, 187, 209
 meaning relations, 226
 nouns, 133, 146, 160
 nouns as nominal pre-modifiers, 107, 116, 132, 144, 148, 165, 170, 181, 209, 222
 meaning relations, 222
 semantic classes, 174
 semantic relations, 178
 novels, 101, 127

 passive voice, 72, 74, 125, 221
 Pearson correlations, 59, 133
 phrasal-prepositional verbs, 48
 popularization, 129
 pre-modifying nouns, 9, 148
 preposition + *-ing* clause, 200

 prepositional phrases, 65, 94, 107, 113, 132, 135, 144, 151, 163, 190, 209–10, 232
 functions, 192
 meaning relations, 237
 specific prepositions, 190
 progressive aspect, 29, 137

 qualitative analysis, 48–9

 register, 37–9, 67–70, 251
 relative clauses, 13, 24–5, 125, 150, 161, 168, 180, 190, 209, 232
 Reppen, 59, 70, 87, 91, 258, 268
 Rissanen, 30, 37, 259, 268
 Rohdenburg, 218, 268
 Rosenbach, 170–1, 268, 271

 Schneider, 33, 36, 259, 269
 Science (journal), 56–7
 science research writing, 50–1, 132
 semi-modals, 46, 137
 shell nouns, 74
 Smith, 29, 178, 266, 269
 speech, 32–6
 stance, 75
 structural elaboration, 16
 structural type, 91
 Swales, 73–5, 268–9
 syntactic function, 92, 141
 Szmrecsanyi, 45, 171–2, 258, 263, 270–1

 T2K-SWAL corpus, 57–8
 textbooks, 101
The Philosophical Transactions of the Royal Society, 56, 126
 Time Magazine, 57–8
 Traugott, 264
 T-unit, 16

 verbs, 154

 Wells, 76, 159
 writing development, 251

