Change in Contemporary English

Based on the systematic analysis of large amounts of computer-readable text, this book shows how the English language has been changing in the recent past, often in unexpected and previously undocumented ways. The study is based on a group of matching corpora, known as the ‘Brown family’ of corpora, supplemented by a range of other corpus materials, both written and spoken, drawn mainly from the later twentieth century. Among the matters receiving particular attention are the influence of American English on British English, the role of the press, the ‘colloquialization’ of written English, and a wide range of grammatical topics, including the modal auxiliaries, progressive, subjunctive, passive, genitive and relative clauses. These subjects build an overall picture of how English grammar is changing, and the linguistic and social factors that are contributing to this process.

Geoffrey Leech is Emeritus Professor of English Linguistics in the Department of Linguistics and English Language at Lancaster University.

Marianne Hundt is Professor of Linguistics in the Department of English at the University of Zürich.

Christian Mair is Professor of English Linguistics in the Department of English at the University of Freiburg.

Nicholas Smith is Lecturer in English Language and Linguistics in the School of English, Sociology, Politics and Contemporary History at the University of Salford.
STUDIES IN ENGLISH LANGUAGE

General editor
Merja Kytö (Uppsala University)

Editorial Board
Bas Aarts (University College London), John Algeo (University of Georgia),
Susan Fitzmaurice (Northern Arizona University), 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:

Christian Mair: Infinitival Complement Clauses in English: A Study of Syntax in Discourse
Charles F. Meyer: Apposition in Contemporary English
Jan Firbas: Functional Sentence Perspective in Written and Spoken Communication
Izchak M. Schlesinger: Cognitive Space and Linguistic Case
Katie Wales: Personal Pronouns in Present-Day English
Laura Wright: The Development of Standard English, 1300–1800: Theories, Descriptions,
Conflicts
Charles F. Meyer: English Corpus Linguistics: Theory and Practice
Stephen J. Nagle and Sara L. Sanders (eds.): English in the Southern United States
Anne Curzan: Gender Shifts in the History of English
Kingsley Bolton: Chinese Englishes
Irma Taavitsainen and Päivi Pahta (eds.): Medical and Scientific Writing in Late
Medieval English
Elizabeth Gordon, Lyle Campbell, Jennifer Hay, Margaret Maclagan, Andrea Sudbury
and Peter Trudgill: New Zealand English: Its Origins and Ecolution
Raymond Hickey (ed.): Legacies of Colonial English
Merja Kytö, Mats Ryden and Erik Smitterberg (eds.): Nineteenth-Century English:
Stability and Change
John Algeo: British or American English? A Handbook of Word and Grammar Patterns
Christian Mair: Twentieth-Century English: History, Variation and Standardization
Evelien Keizer: The English Noun Phrase: The Nature of Linguistic Categorization
Raymond Hickey: Irish English: History and Present-Day Forms
Günter Rohdenburg and Julia Schlüter (eds.): One Language, Two Grammars?:
Differences between British and American English
Laurel J. Brinton: The Comment Clause in English
Lieselotte Anderwald: The Morphology of English Dialects: Verb Formation in
Non-standard English
Contents

List of figures  page x
List of tables  xiv
Preface  xix
Abbreviations and symbolic conventions  xxiv

1 Introduction: ‘grammar blindness’ in the recent history of English?  1
  1.1 Grammar is more than an arbitrary list of shibboleths  1
  1.2 Grammatical changes: proceeding slowly and invisible at close range?  7
  1.3 A frame of orientation: previous research on recent and ongoing grammatical changes in English  16
  1.4 Conclusion  22

2 Comparative corpus linguistics: the methodological basis of this book  24
  2.1 (Computer) corpus linguistics: the Brown Corpus and after  24
  2.2 Comparable corpora and comparative corpus linguistics  27
  2.3 The methodological basis of comparable corpus linguistics  31
  2.4 Stages of investigation  33
    (A) Rationalize the mark-up of the corpora  33
    (B) Undertake annotation of the corpora  33
    (C) Use search and retrieval software to identify and extract recurrent formal features in the corpus  34
    (D) Refine the comparative analysis  34
    (D1) Derive difference-of-frequency tables  35
    (D2) Derive difference-of-frequency tables from inter-corpus comparisons  36
vi Contents

(D3) Undertake further categorization of instances of features found in the corpora 36
(E) Further qualitative analysis, examining individual instances, or clusters of instances, in both corpora 37
(F) Functional interpretation of findings 37

2.5 Further details and explanations of the stages of investigation 37
2.5.1 (B) Annotation 37
2.5.2 (C) Search expressions in CQP 38
2.5.3 (D1) Frequency across genres and subcorpora 40
2.5.4 (D2) External comparisons 43
2.5.5 (D3) Further categorization of instances found in the corpora 45
2.5.6 (E) Further qualitative analysis 47
2.5.7 (F) Functional interpretation of findings on all levels 49

2.6 Conclusion 50

3 The subjunctive mood 51
3.1 Introduction 51
3.2 The revival of the mandative subjunctive 52
3.2.1 Overall developments of the mandative subjunctive 53
3.2.2 Is the mandative subjunctive losing its formal connotations? 57
3.3 The were-subjunctive 61
3.3.1 The were-subjunctive: diachronic development 64
3.3.2 The were-subjunctive: a recessive formal option? 66
3.4 Revival and demise of the subjunctive? An attempt at reconciling apparently contradictory developments 67
3.5 Summary and conclusion 69

4 The modal auxiliaries 71
4.1 The declining use of the modal auxiliaries in written standard English 1961–1991/2 71
4.2 The changing use of the modals in different genres and subcorpora 73
4.3 The changing use of the modals in spoken vs written corpora 76
4.4 The core modals and competing expressions of modality 78
## Contents

4.5 Shrinking usage of particular modals: a more detailed examination  
4.5.1 The modals at the bottom of the frequency list: shall, ought to and need(n’t)  
4.5.2 The semantics of modal decline: may, must and should  
4.6 Conclusion  

5 The so-called semi-modals  
5.1 Auxiliary–lexical verb gradience  
5.2 Overall changes in frequency of semi-modals  
5.3 Further evidence for grammaticalization? Phonetics and semantics  
5.3.1 Phonetic reduction and coalescence: gonna, gotta and wanna  
5.3.2 Signs of abstraction and generalization (semantic weakening)  
5.4 The ecology of obligation/necessity  
5.5 Conclusion  

6 The progressive  
6.1 Introduction  
6.2 Basic and special uses of the progressive  
6.3 Historical background  
6.4 Overview of recent distribution patterns  
6.4.1 Distribution in written BrE and AmE  
6.4.2 Distribution in contemporaneous BrE speech and other registers  
6.5 Present progressive active  
6.5.1 Quotations and contracted forms  
6.5.2 Stative verbs  
6.5.3 Subject type and reference  
6.5.4 Special uses  
6.6 The progressive passive  
6.7 The progressive in combination with modal auxiliaries  
6.7.1 Modal auxiliary + be -ing  
6.7.2 Will + be -ing  
6.8 Summary and conclusion  

7 The passive voice  
7.1 Introduction  
7.2 The be-passive  
7.3 The get-passive  
7.4 The mediopassive  
7.5 Summary and conclusion
viii Contents

8 Take or have a look at a corpus? Expanded predicates in British and American English 166
  8.1 The state of the art 167
  8.2 Hypotheses 170
  8.3 Defining the variable 173
  8.4 Results 174
    8.4.1 Stylistic variation 174
    8.4.2 Diachronic variation 175
    8.4.3 Regional variation 175
  8.5 Summary 179

9 Non-finite clauses 181
  9.1 Introduction: long-term trends in the evolution of English non-finite clauses 181
  9.2 Changes in non-finite clauses I: case studies of individual matrix verbs 186
    9.2.1 Help + infinitive 187
    9.2.2 Prevent/stop + NP + (from) + gerund 193
    9.2.3 Start and stop in catenative uses 195
    9.2.4 Want to 199
    9.2.5 Assessing the speed of changes 201
  9.3 Changes in non-finite clauses II: statistical trends in the tagged corpora 201
  9.4 Conclusion 204

10 The noun phrase 206
  10.1 Parts of speech: an overall survey 207
  10.2 Nouns and noun sequences 211
    10.2.1 Common nouns 212
    10.2.2 Proper nouns, including proper nouns as acronyms 212
  10.3 Noun sequences and other juxtapositions 214
    10.3.1 Noun + common noun sequences 216
    10.3.2 Noun sequences with plural attributive nouns 219
    10.3.3 Sequences of proper nouns 221
  10.4 The s-genitive and the of-genitive 222
    10.4.1 The s-genitive 223
    10.4.2 The of-genitive 224
  10.5 Relative clauses 226
    10.5.1 Wh- relative clauses 228
    10.5.2 That relative clauses 229
    10.5.3 Zero relative clauses 231
    10.5.4 Pied-piping vs preposition stranding 231
  10.6 Summary and conclusion 233
11 Linguistic and other determinants of change 236
  11.1 The functional and social processes of change 236
  11.2 Grammaticalization 237
  11.3 Colloquialization
    11.3.1 Contracted negatives and verb forms 240
    11.3.2 Not-negation vs no-negation 241
    11.3.3 Questions 242
    11.3.4 Other plausible grammatical signs of colloquialization 243
    11.3.5 Punctuation 244
    11.3.6 Problems and issues concerning colloquialization 245
  11.4 Densification of content 249
  11.5 Americanization? 252
    11.5.1 ‘Americanization’ in relation to other trends 256
    11.5.2 ‘Americanization’ and sociolinguistic globalization 258
  11.6 Other trends 259
    11.6.1 Democratization: ironing out differences 259
    11.6.2 Language prescriptions 263
    11.6.3 Analyticization? 264
  11.7 Conclusion 267

Appendix I The composition of the Brown Corpus 273
Appendix II The C8 tagset used for part-of-speech tagging of the four corpora 276
Appendix III Additional statistical tables and charts 281
References 314
Index 335
Figures

Figure 0.1 The four matching corpora on which this book focuses

Figure 1.1 Matching one-million-word corpora of written English

Figure 1.2 Advertisement for Fels-Naptha Soap, *The Times*, 31 December 1915, page 15

Figure 2.1 Comparison of the use of the passive in LOB and F-LOB Corpora

Figure 2.2 A fragment of an annotated database of the progressive passive, using Excel

Figure 3.1 *Should*-periphrasis vs mandative subjunctive in written AmE and BrE

Figure 3.2 Indicative, *should*-periphrasis and subjunctive after mandative expressions in ICE-GB (frequency per million words)

Figure 3.3 Subjunctive *were* vs indicative *was* in hypothetical/unreal conditional constructions

Figure 4.1 Frequencies of modals in the four written corpora: comparing 1961 with 1991/2

Figure 4.2 Modal auxiliaries in American English, 1961–92

Figure 4.3 Modal auxiliaries in British English, 1961–91

Figure 4.4 Comparison of DSEU and DICE: modals in spoken BrE in the 1960s and the early 1990s

Figure 4.5 *May* – change in frequency of senses (analysis of every third example) in the Brown family of corpora

Figure 4.6 *Should* – change in frequency of senses in the Brown family of corpora

Figure 4.7 *Must* – change in frequency of senses (analysis of every third example) in the Brown family of corpora

Figure 5.1 The auxiliary–main verb gradient, from Quirk *et al.* (1985: 137)

Figure 5.2 Change of frequency in the semi-modals in written English (the Brown family, AmE and BrE combined)
List of figures xi

Figure 5.3 Frequency of semi-modals in spoken British English: increase in use based on the comparison of the DSEU and DICE mini-corpora 99

Figure 5.4 An 'apparent-time' study: comparison of age groups of speakers in the BNC demographic subcorpus (BNCdemog): distribution of modals and semi-modals according to age 103

Figure 5.5 A study in apparent time: contracted forms gonna, gotta and wanna, as percentage of full and contracted forms, in the spoken BNC (based on Krug 2000: 175) 106

Figure 5.6 Meanings of can and BE able to 112

Figure 6.1 Distribution of the progressive in ARCHER (based on Hundt 2004a: 69) 122

Figure 6.2 Progressives by subcorpora in LOB and F-LOB (1961–1991): changes in frequency pmw 123

Figure 6.3 Progressives by subcorpora in Brown and Frown (1961–1992): frequencies pmw 123

Figure 6.4 Distribution of the progressive in genres of the full ICE-GB corpus (1990–92): frequencies pmw 125

Figure 6.5 Progressives by broad genre category in the DSEU (1958–69) and DICE (1990–92): frequencies pmw 126

Figure 6.6 Present progressive passive in LOB and F-LOB: frequencies pmw 138

Figure 6.7 Past progressive passive in LOB and F-LOB: frequencies pmw 138

Figure 6.8 Non-progressive present passive in LOB and F-LOB: frequencies pmw 138

Figure 6.9 Non-progressive past passive in LOB and F-LOB: frequencies pmw 138

Figure 7.1 Finite non–progressive be–passives in the Brown family of corpora: frequencies per million words 149

Figure 7.2 Get–passives (all forms) in the Brown family of corpora: frequencies per million words 156

Figure 7.3 Semantics of the get–passive in the Brown family of corpora (based on pooled frequencies for the two subperiods) 157

Figure 8.1 Expanded predicates across different text types (frequencies pmw) 174

Figure 8.2 Diachronic development of light verbs in expanded predicates (proportion of light verbs per number of expanded predicates) 176

Figure 8.3 Expanded predicates in spoken British and American English 177
### List of figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.4</td>
<td>Expanded predicates with variable use of <em>have</em> and <em>take</em> in spoken British and American English (relative frequencies)</td>
<td>177</td>
</tr>
<tr>
<td>8.5</td>
<td>Expanded predicates in written and spoken English (pmw)</td>
<td>179</td>
</tr>
<tr>
<td>9.1</td>
<td><em>To</em>- vs bare infinitives with <em>help</em> (all construction types) in five corpora – diachronic trends</td>
<td>189</td>
</tr>
<tr>
<td>9.2</td>
<td>Infinitival and gerundial complements with <em>start</em> in five corpora – diachronic trends</td>
<td>196</td>
</tr>
<tr>
<td>9.3</td>
<td>Gerundial complements with <em>start</em> and <em>stop</em> in five corpora – diachronic trends</td>
<td>197</td>
</tr>
<tr>
<td>9.4</td>
<td>Infinitival and gerundial complements with <em>begin</em> and <em>start</em> in two spoken corpora – regional comparison between British (BNCdemog) and American English (ANC)</td>
<td>198</td>
</tr>
<tr>
<td>10.1</td>
<td>Increase of noun + common noun sequences in AmE (Brown to Frown) and BrE (LOB to F-LOB). Frequencies normalized to pmw</td>
<td>217</td>
</tr>
<tr>
<td>10.2</td>
<td>Increase in plural attributive nouns in N+N sequences</td>
<td>221</td>
</tr>
<tr>
<td>10.3</td>
<td>Increase in frequency of proper noun + proper noun sequences 1961–1991/2 in Brown, Frown, LOB and F-LOB</td>
<td>222</td>
</tr>
<tr>
<td>10.4</td>
<td>Increase in frequency of <em>s</em>-genitives 1961–1991/2 in Brown, Frown, LOB and F-LOB (frequencies pmw)</td>
<td>223</td>
</tr>
<tr>
<td>10.5</td>
<td>Change of frequency of the <em>of</em>-genitive in relation to the <em>s</em>-genitive between 1961 and 1991/2, expressed as a percentage of all ‘genitives’</td>
<td>225</td>
</tr>
<tr>
<td>10.6</td>
<td>Change of frequency of the three types of relativization 1961–1991/2: decline of the <em>wh</em>– relatives, and increasing frequency of the <em>that</em>– and zero relatives. Expressed as a percentage of all (finite) relative clauses apart from those with an adverbial gap.</td>
<td>227</td>
</tr>
<tr>
<td>10.7</td>
<td>Increasing use of <em>that</em>-relative clauses 1961–1991/2 in AmE (Brown → Frown) and BrE (LOB → F-LOB): frequencies pmw</td>
<td>229</td>
</tr>
<tr>
<td>10.8</td>
<td>A small increase in preposition stranding in relative clauses in the three varieties of relative clause between 1961 and 1991 in written BrE.</td>
<td>232</td>
</tr>
<tr>
<td>11.1</td>
<td>Abstract nominalizations in AmE: frequencies pmw</td>
<td>250</td>
</tr>
<tr>
<td>11.2</td>
<td>Abstract nominalizations in BrE: frequencies pmw</td>
<td>251</td>
</tr>
<tr>
<td>11.3</td>
<td>A follow-my-leader pattern: declining frequency of the core modals in AmE and in BrE.</td>
<td>254</td>
</tr>
<tr>
<td>11.4</td>
<td>Increasing use of contractions in AmE and BrE: summary (frequencies pmw)</td>
<td>257</td>
</tr>
<tr>
<td>11.5</td>
<td>Decline of titular nouns preceding personal names in AmE (frequencies pmw)</td>
<td>260</td>
</tr>
</tbody>
</table>
List of figures  xiii

Figure 11.6  Decline of titular nouns preceding personal names in BrE (frequencies pmw)  260
Figure 11.7  Periphrastic comparatives as a percentage of all comparative forms  266

List of Figures in Appendix III
Figure A6.1  Distribution of present progressives (active) in LOB and F-LOB across subcorpora: frequencies pmw  296
Figure A6.2  Distribution of present progressives (active) in Brown and Frown across subcorpora: frequencies pmw  296
Figure A6.3  Distribution of present progressives (active) in 1961, Brown versus LOB: frequencies pmw  297
Figure A6.4  Distribution of present progressives (active) in 1991/2, Frown versus F-LOB: frequencies pmw  297
# Tables

<p>| Table 1.1 | Whom (interrogative and relative function) in four matching corpora | page 13 |
| Table 2.1 | Brown and LOB Corpora compared in terms of genres, number of texts and number of words | 26 |
| Table 2.2 | Comparisons between corpora in the Brown family | 29 |
| Table 2.3 | Comparison of use of the passive in the LOB and F-LOB corpora | 35 |
| Table 2.4 | A partial repetition of Table 2.3: passives | 40 |
| Table 2.5 | Another partial repetition of Table 2.3: passives | 41 |
| Table 2.6 | First-person singular pronouns (I, me) in the Brown and Frown Corpora | 42 |
| Table 3.1 | Distribution of mandative subjunctives across text categories (figures in brackets give the frequency per million words) | 58 |
| Table 3.2 | Mandative subjunctives and periphrastic constructions: active vs passive VPs | 59 |
| Table 3.3 | Mandative subjunctives and periphrastic constructions in written and spoken English (percentages are given for subjunctive only) | 60 |
| Table 3.4 | Subjunctive were vs indicative was in hypothetical/unreal conditional constructions | 65 |
| Table 3.5 | Distribution of were-subjunctives across text types (figures per million words are given in brackets) | 67 |
| Table 4.1 | Change of frequency of the core modals in subcorpora | 75 |
| Table 5.1 | Evolving rivalry between must and HAVE to in terms of frequency | 99 |
| Table 5.2 | Approximate frequency count of modals and semi-modal in the LCSAE | 100 |
| Table 5.3 | Frequency of modals and semi-modal in the demographic subcorpus BNCdemog: the conversational part of the BNC | 102 |</p>
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1</td>
<td>Expanded predicates with variable use of <em>have</em> and <em>take</em> in spoken British and American English (raw frequencies)</td>
<td>178</td>
</tr>
<tr>
<td>9.1</td>
<td><em>Help</em> + infinitive in spoken American and British English</td>
<td>191</td>
</tr>
<tr>
<td>9.2</td>
<td><em>Help</em> in the Diachronic Corpus of Present-Day Spoken English</td>
<td>192</td>
</tr>
<tr>
<td>9.3</td>
<td><em>Prevent</em> NP from V-ing vs <em>prevent</em> NP V-ing in five corpora</td>
<td>194</td>
</tr>
<tr>
<td>9.4</td>
<td><em>Stop</em> NP from V-ing vs <em>stop</em> NP V-ing in four corpora</td>
<td>194</td>
</tr>
<tr>
<td>9.5</td>
<td><em>Stop/start</em> + gerund by speaker age in the spoken-demographic BNC</td>
<td>199</td>
</tr>
<tr>
<td>9.6</td>
<td><em>Start</em> + gerund in DCPSE</td>
<td>199</td>
</tr>
<tr>
<td>9.7</td>
<td><em>Stop</em> + gerund in DCPSE</td>
<td>199</td>
</tr>
<tr>
<td>9.8</td>
<td><em>Want</em> in five corpora</td>
<td>200</td>
</tr>
<tr>
<td>9.9</td>
<td><em>To</em>-infinitives in four corpora</td>
<td>202</td>
</tr>
<tr>
<td>9.10</td>
<td><em>To</em>-infinitives as percentages of all verbal tags in four corpora</td>
<td>202</td>
</tr>
<tr>
<td>9.11</td>
<td>Prepositional gerunds in four corpora</td>
<td>203</td>
</tr>
<tr>
<td>10.1</td>
<td>Percentage changes in the frequency of part-of-speech categories</td>
<td>207</td>
</tr>
<tr>
<td>10.2</td>
<td>Frequency of nouns in the LOB and F-LOB corpora, showing major genre subdivisions of the corpora</td>
<td>211</td>
</tr>
<tr>
<td>10.3</td>
<td>Increasing frequency of various subcategories of noun in AmE and BrE</td>
<td>212</td>
</tr>
<tr>
<td>10.4</td>
<td>Change in relative frequency of subcategories of proper noun in AmE and BrE, based on 2% randomized samples of the four corpora</td>
<td>213</td>
</tr>
<tr>
<td>10.5</td>
<td>Expressions referring to the President of the United States in Brown and Frown</td>
<td>214</td>
</tr>
<tr>
<td>10.6</td>
<td>Additional noun + common noun sequences as a percentage of additional nouns in the 1991/2 corpora</td>
<td>219</td>
</tr>
<tr>
<td>11.1</td>
<td><em>Not</em>-negation and <em>no</em>-negation in AmE (Brown, Frown) and BrE (LOB, F-LOB)</td>
<td>242</td>
</tr>
<tr>
<td>11.2</td>
<td>Questions in AmE and BrE</td>
<td>243</td>
</tr>
<tr>
<td>11.3</td>
<td>Decreasing use of main verb <em>have</em> constructed as an auxiliary, and increasing use of <em>do</em>-support with <em>have</em> in negation and inversion</td>
<td>255</td>
</tr>
<tr>
<td>11.4</td>
<td>Decline of gender-neutral <em>he</em> and rise of alternative pronominal expressions</td>
<td>262</td>
</tr>
<tr>
<td>11.5</td>
<td>Periphrastic and inflectional comparison in AmE</td>
<td>265</td>
</tr>
<tr>
<td>11.6</td>
<td>Periphrastic and inflectional comparison in BrE</td>
<td>265</td>
</tr>
<tr>
<td>11.7</td>
<td>Number of adjectives exhibiting both inflectional and periphrastic comparison in the same corpus</td>
<td>267</td>
</tr>
</tbody>
</table>
xvi List of tables

Table 11.8. Summary table: postulated explanatory trends, together with the increases and decreases of frequency they help to explain 271

List of Tables in Appendix III 281
Table A3.1. Subjunctive vs should-periphrasis in four parallel corpora 281
Table A3.2. Indicative, should-periphrasis and subjunctive after mandative expressions in ICE-GB 282
Table A4.1. Frequencies of modals in the four written corpora: comparing 1961 with 1991 283
Table A4.2. Modal auxiliaries in AmE and BrE respectively 283
Table A4.3 Comparison of DSEU and DICE: modals in spoken BrE in the 1960s and the early 1990s 284
Table A4.4. May – change in frequency of senses (analysis of every third example) 284
Table A4.5. Should – change in frequency of senses 284
Table A4.6. Must – change in frequency of senses (analysis of every third example) 285
Table A4.7. May, should and must – changes in frequency of senses in spoken mini-corpora 285
Table A5.1. Frequencies of semi-modals in the Brown family of corpora 286
Table A5.2. Changing frequencies of semi-modals in the two British spoken mini-corpora: DSEU (1958–69) and DICE (1990–92) 287
Table A5.3. A study in apparent time: modals and semi-modals in the BNC demographic subcorpus 287
Table A6.1. Distribution of progressives across the paradigm in LOB, F-LOB, Brown and Frown (whole corpus frequencies) 288
Table A6.2. Frequencies of progressives relative to estimated count of non-progressives in LOB and F-LOB 289
Table A6.3. Distribution of all progressives in written AmE (1961–1991/92) 289
Table A6.4. Genre distribution of progressives in spoken BrE: DSEU (1958–69) and DICE (1990–92) 289
Table A6.5. Distribution of present progressives (active) outside quotations, in LOB and F-LOB 290
Table A6.6. Contracted forms of present progressive (active) in LOB and F-LOB 290
Table A6.7. Contracted forms of present progressive (active) outside quotations in LOB and F-LOB 291
List of tables  xvii

Table A6.8. Contracted forms in all syntactic environments in LOB and F-LOB  291
Table A6.9. Distribution of the present progressive (active) of verbs lending themselves to stative interpretation in LOB, F-LOB, Brown and Frown  292
Table A6.10. Subject person and number of present progressives (active) in LOB and F-LOB  293
Table A6.11. Futurate use of present progressive (active) in LOB and F-LOB: clear cases only  294
Table A6.12. Futurate use of present progressive (active) in Brown and Frown: clear cases only (based on a 1 in 4 sample)  294
Table A6.13. Constructions referring to the future in corpora of recent British English (LOB and F-LOB): raw and proportional frequencies  294
Table A6.14. Frequencies of modal + be -ing and modal + infinitive constructions in LOB and F-LOB: whole corpus frequencies  295
Table A6.15. Modal + be -ing and modal + infinitive in Brown/Frown: whole corpus frequencies  295
Table A6.16. Distribution of interpretive use of the progressive (present tense), in LOB and F-LOB, based on clearest cases  296
Table A6.17. Functions of will + be -ing: estimated frequencies in LOB and F-LOB  296
Table A7.1. Finite non-progressive be-passives in the Brown family of corpora  297
Table A7.2. Composition of the DCPSE subcorpus  298
Table A7.3. Get-passives in the Brown family of corpora  298
Table A8.1. Retrieved expanded predicates (types) in the Brown family of corpora  299
Table A8.2. Expanded predicates across text types in the Brown family of corpora  300
Table A9.1. To- vs bare infinitives with help (all construction types) in five corpora  301
Table A9.2. To-inf. vs V-ing after start  301
Table A9.3. Begin/start + infinitive by speaker age in the spoken-demographic BNC  301
Table A9.4. Start/stop + gerund in five corpora  302
Table A9.5. To-inf. vs -ing after begin and start in the spoken-demographic BNC and the spoken ANC  302
Table A10.1a. Comparison of tag frequencies in LOB and F-LOB: change in the frequency of parts of speech in BrE 1961–91  302
xviii  List of tables

Table A10.1b.  Comparison of tag frequencies in Brown and Frown: change in the frequency of parts of speech in AmE, 1961–92  

Table A10.2.  Frequency of selected noun subcategories in the LOB and F-LOB Corpora  

Table A10.3a.  Noun combinations in the language of the Press (A–C)  

Table A10.3b.  Noun combinations in the language of Learned writing (J)  

Table A10.4a.  Increase of noun + common noun sequences between LOB and F-LOB  

Table A10.4b.  Development of noun + common noun sequences between Brown and Frown  

Table A10.5.  Plural attributive nouns in AmE and BrE  

Table A10.6.  Frequency of proper noun + proper noun sequences  

Table A10.7a.  S-genitives in American English: Brown vs Frown  

Table A10.7b.  S-genitives in British English: LOB vs F-LOB  

Table A10.8.  Of-genitives in AmE (Brown and Frown) and in BrE (LOB and F-LOB): a sample from 2% of all of-phrases  

Table A10.9a.  Decreasing use of wh- relative pronouns (who, whom, whose, which) in AmE (Brown and Frown)  

Table A10.9b.  Decreasing use of wh- relative pronouns (who, whom, whose, which) in BrE (LOB and F-LOB)  

Table A10.10.  The relative pronoun which in AmE (Brown and Frown) and in BrE (LOB and F-LOB)  

Table A10.11a.  Increasing use of that-relative clauses in AmE (Brown and Frown)  

Table A10.11b.  Increasing use of that-relative clauses in BrE (LOB and F-LOB)  

Table A10.12a.  Zero relative clauses in AmE (Brown and Frown)  

Table A10.12b.  Zero relative clauses in BrE (LOB and F-LOB)  

Table A10.13.  Pied-piping in wh- relative clauses (LOB and F-LOB)  

Table A11.1.  Lexical density in AmE and BrE (1961–1991/2)  

Table A11.2a.  Punctuation marks: a comparison of AmE and BrE changes 1961–91  

Table A11.2b.  Some punctuation marks: a comparison of B-LOB (BrE 1931) and F-LOB (BrE 1991)  

Table A11.3.  Abstract nominalizations in AmE  

Table A11.4.  Abstract nominalizations in BrE  

Table A11.5.  Decline of titular nouns preceding personal names: raw frequencies
Preface

This book aims to give an account of how the English language has been changing recently, focusing especially on (a) the late twentieth century, (b) the written standard language, (c) American and British English, (d) grammatical rather than lexical change, and using the empirical evidence of computer corpora.

Corpus linguistics is now a mainstream paradigm in the study of languages, and the study of English in particular has advanced immeasurably through the availability of increasingly rich and varied corpus resources. This applies to both synchronic and diachronic research. However, this book presents, we argue, a new kind of corpus-based historical research, with a narrower, more intense focus than most, revealing through its rather rigorous methodology how the language (more especially the written language) has been developing over a precisely defined period of time in the recent past.

The period on which the book concentrates is the thirty years between the early 1960s and the early 1990s, and the four corpora that it studies in most detail are those which go increasingly by the name of the ‘Brown family’: the Brown corpus (American English, 1961); the Lancaster–Oslo/Bergen corpus (British English, 1961); the Freiburg–Brown corpus (American English, 1992); and the Freiburg–Lancaster–Oslo/Bergen corpus (British English, 1991).¹

These corpora, described in more detail in Chapter 2 (section 2.2) and in Appendix II, are reasonably well known, and have been studied as a group, not only by ourselves, but by others, since the completion of this corpus quartet in the mid-1990s. All four corpora are available to researchers around the world, and can be obtained under licence from either ICAME at the Aksis centre, University of Bergen, or the Oxford Text Archive, University of Oxford.² However, we venture to claim that as authors of this book we have been more intimately engaged with these corpora than any other research group: in their compilation, their annotation and their analysis. Indeed, this

¹ An informative manual of information for the Brown family of corpora, including their POS tagging, is provided by Hinrichs et al. (forthcoming).
² The web addresses of these two corpus resource agencies are as follows: http://icame.uib.no/ and http://ota.abds.ac.uk/.
intimacy entitles us to feel a certain familial affection for these textual time-capsules, and almost invariably (like many others) we refer to them by their acronymic nicknames: Brown, LOB, Frown and F-LOB.  

The strength of these four corpora lies in their comparability: the fact that they are constructed according to the same design, having virtually the same size and the same selection of texts and genres represented by 500 matching text samples of c. 2,000 words. This means that we can use the Brown family as a precision tool for tracking the differences between written English in 1961 and in 1991/2. How has the English language changed, in these two leading regional varieties, over this thirty-year generation gap? The findings brought to light by this comparison between matching corpora are fascinating: they reveal, for the first time, or at least with a new sense of accuracy, how significant are the changes in a language that take place over even such a short timespan of thirty years. Even though these changes, as we report them, are almost entirely matters of changing frequency of use, they often show a high degree of statistical significance.  

The affection we feel for this corpus family does not blind us to their considerable limitations (see section 2.1), notably their restriction to the standard written language. We have therefore taken care to supplement the evidence they provide with analyses of other corpora relating to the later twentieth century, so as to enlarge and corroborate our findings on how the language has recently been changing. In extending our range in this way, most

---

3 The explanations of these names for corpora, as well as other abbreviations, are found in the list of 'Abbreviations and symbolic conventions', pp. xxvii–xxx.

4 Significance levels are shown, where appropriate, by asterisks: *, **, *** in the quantitative tables – see the table of Abbreviations and symbolic conventions.
important have been the corpora that record indications of what has been happening to the spoken language. The Diachronic Corpus of Present-Day Spoken English (DCPSE), released in 2006, has made it possible to study, over the sample period of time, changes in the spoken language, though not under the same rigorous conditions of comparability that apply to the Brown family. In addition, the British National Corpus (BNC), though it has no reliable diachronic dimension, gives us a large (ten-million-word) well-sampled subcorpus of spoken English from the early 1990s. Both of these corpora are limited to British English: but we have been able to consult the CIC (Cambridge International Corpus) and LCSAE (Longman Corpus of Spoken American English, comparable in date and method of collection to the spoken demographic subcorpus of the BNC) to see how that presumably most trail-blazing variety of the language – spoken American English – compares with others. Again, there is only an indirect diachronic dimension here, through the study of ‘apparent time’ by comparison of different age groups of speakers. But at least we are able to speculate on tangible evidence about how the spoken American variety has been moving in the period under review.

Apart from these (necessarily imperfect and incomplete) comparisons between corpora of speech and writing, we have also been able to extend our range, when need arises, along the diachronic dimension. In the months preceding the publication of this book, we were able to make limited use of the newest member of the Brown family – though oldest in date – the Lancaster 1931 Corpus (inevitably nicknamed ‘B-LOB’ for ‘before-LOB’), sampled from a seven-year period centring on 1931, and so effectively providing us with three equidistant reference points, 1931 (± 3 years), 1961 and 1991/2, for further diachronic comparison. For even greater historical depth, we have occasionally used the ARCHER corpus and the OED citation bank. These valuable resources again lack the strict comparability criterion of the Brown

---

5 The DCPSE, consisting of 885,436 words, and compiled by Bas Aarts and associates at the Survey of English Usage, University College London, consists of transcribed British spoken texts originally collected as parts of two different corpora: (a) the Survey of English Usage corpus (of which the spoken part was later largely incorporated into the London–Lund Corpus) collected in 1958–1978; and (b) the ICE-GB corpus collected in 1990–1992. Geoffrey Leech is grateful to Bas Aarts for letting him have an advance copy of DCPSE at a point when it was timely for drafting certain chapters of this book.

6 It should be mentioned that there are several corpora of present-day spoken English of which we have not made detailed use, since, although admirable for other types of research, they are either too small for our present purposes (e.g. the Santa Barbara Corpus of Spoken American English) or too genre-restricted (e.g. MICASE, Corpus of Spoken Professional American English, the Switchboard corpus).

7 This corpus, now in a provisional pre-release form, has been compiled by Nicholas Smith, Paul Rayson and Geoffrey Leech with the financial support of the Leverhulme Foundation. With further support from the Leverhulme Foundation, we will shortly have yet another member of the Brown family, with a corpus of BrE at the beginning of the twentieth century (1901 ± 3 years to be precise). However, this corpus, provisionally called Lanc-1901, was not completed in time for its results to be used in this book.
family, but allow corpus-based investigations of trends going back to EModE (in the case of ARCHER) and to OE (in the case of the OED citation bank). Turning towards the future: we have not been able to draw on more recent progeny of the Brown family, since none are yet available; but the ‘corpus of last resort’ these days, the World Wide Web (see a number of contributions to Hundt et al. 2007), has sometimes given us persuasive evidence about what has been happening since the early 1990s.8

What has become obvious is that the corpus resources available for recent diachronic research do not comprise a static platform for research, but a moving staircase: every year new text resources become available, in increasing numbers and increasing size, enhancing our evidential basis for researching the recent development of the language. In such a situation of continuing advance, it is a reasonable compromise to adopt the position we have taken – to focus on the four tried-and-tested Brown family corpora, while using other corpora where it is particularly rewarding or important (as well as feasible) to do so.

The unavoidable assumption of incompleteness is familiar in many fields of scientific endeavour: if researchers before publication waited until complete results and complete answers were available, there would be no publication. Certainly, it would have been easy for us to engage in further research on the range of topics we have investigated here, collecting or consulting further corpora, carrying out deeper analyses, and so on, without reaching a natural endpoint. We hope that in spite of its existing limitations, this book will be felt to have achieved a valuable conspectus of new or recent findings across a wide variety of grammatical topics. Although we have taken care to achieve a consistent perspective and framework of research throughout the book, readers may notice some lack of consistency in the kinds of coverage of corpus analyses offered in individual chapters. In the ‘moving staircase’ scenario described above, this is almost inevitable, and there is after all no harm in a book which reflects to some extent the different emphases, interests and strengths of individual chapter authors.

One of the most positive achievements of our collaboration is the uniform part-of-speech annotation (or POS tagging) of all four corpora – all five, if one includes the 1931 corpus. We have used the same software annotation practices (the Lancaster tagger CLAWS, the supplementary tagger Template Tagger and the enriched C8 tagset of grammatical categories – see Appendix II and also the detailed tagging guide in Hinrichs et al. forthcoming). This has enabled the corpora to be compared, grammatically, on an equal footing, using equivalent search and retrieval patterns to extract instances of abstract constructions, such as progressives, and in some important instances (e.g.

8 Paul Baker of Lancaster University has provisionally compiled a twenty-first century web-derived corpus on the Brown model, and this will eventually take its place in the Brown family of corpora.
zero relativization) even grammatical categories not explicitly realized in surface structure. Here again, however, we have not managed to achieve complete consistency of treatment: the three corpora LOB, F-LOB and Frown have all been manually post-edited after automatic tagging, while the Brown corpus, the earliest of all to be compiled and tagged, has not undergone the manual post-edit with the new set of tags. This has meant a lower degree of confidence (with an initial error margin of c. 2 per cent) in the correctness of some results in the American English (AmE) comparison of Brown and Frown, alongside the more accurate British English (BrE) comparison of LOB and F-LOB. However, this margin of error has been minimized by employing a corrective coefficient based on the tagger’s error rates observed in the comparison of pre-corrected and post-corrected versions of the Frown Corpus – see further p. 24, footnote 27. The dictum that ‘Most corpus findings are approximations’ (see section 2.3) is particularly to be taken to heart in interpreting our findings for grammatical constructions and categories in AmE, and this has sometimes led us to give more attention to the results for BrE than those for AmE.

Given that the book focuses on changes in grammar, the POS tagging combined with powerful CQP search software (see section 2.4 C) has enabled us, without aiming at comprehensiveness, to achieve a broad grammatical coverage of the language. After two introductory chapters, the next seven chapters concentrate on topics relating to the verb phrase. They cover the subjunctive (Chapter 3), the modal auxiliaries (Chapter 4), the so-called semi-modals (Chapter 5), the progressive aspect (Chapter 6), the passive (Chapter 7), expanded predicates such as have/take a look (Chapter 8) and non-finite constructions (Chapter 9). In Chapter 10 we move on to the noun phrase, enquiring particularly into noun–noun sequences, genitives and relative clauses. In the last chapter, Chapter 11, we seek a synthesis, dealing with social and linguistic determinants of the short-term changes demonstrated in earlier chapters, and extending the book’s coverage by illustrating these determinants with a number of additional linguistic trends.

The book abounds with statistical tables and charts, comparing frequencies (often normalized to occurrences per million words) according to period of

---

Footnotes:

9 Tables and figures relying on approximations based on adjusted automatic tagging counts in this way occur mostly in Chapters 10 and 11, or in the part of Appendix III relating to these chapters. Such tables and figures are indicated by a warning note ‘(automatic)’ or ‘(AmE automatic)’ beneath the relevant table or figure.

10 A simple and obvious point has to be made here: we have naturally given primary attention to areas of English grammar known or suspected to be undergoing change. (In some cases the ‘knowledge’ or ‘suspicion’ comes from our own exploratory study of the corpora.) There are, however, interesting areas of contemporary English grammar that we have not dealt with: for example, we will have nothing to say about corpus findings relating to the choice of singular or plural verb after a collective-noun subject (The team is/are . . . – a construction that has, however, been more than adequately studied elsewhere – see Levin 2001, 2006; Depraetere 2003; Hundt 2006). Our failure to treat a particular topic is not a reliable signal of its lack of interest from the present-day diachronic viewpoint.
time (mostly 1961 vs 1991), region, genre, etc. We have aimed to provide sound corpus description, using inferential statistics to generalize beyond corpus observations, looking at single dependent variables at a time, and interpreting the findings in the framework of a reasonable and robust usage-based model of language change. To avoid cluttering up the descriptive chapters (Chapters 3–11) with statistical details that might obscure the main findings and lines of argument, we have consigned many of the statistical tables and diagrams, particularly the more complex ones, to Appendix III.

The four authors are jointly responsible for the whole work in its final form; nevertheless, it may be of interest to know which authors took particular responsibility for which chapters. They are here identified by their initials: GL: 2, 4, 5, 10, 11; MH: 3, 7, 8, also the References; CM: 1, 9; NS: 6, Appendices. It should be added, however, that the relative input of individual authors can by no means be measured in this way.

This is the appropriate point to acknowledge gratefully our debt to those who helped us in various ways; to Merja Kytö as series General Editor, and to Helen Barton, editor at Cambridge University Press, we owe a great deal for their encouragement, support and forbearance. We also owe much to the research assistants who helped us in the processing of textual data: Lars Hinrichs, Barbara Klein, Luminița-Irinel Trașcă and Birgit Waibel in Freiburg; and Martin Schendzielorz in Heidelberg. We are grateful, too, to Paul Rayson and Sebastian Hoffmann, colleagues at Lancaster; to Gunnel Tottie, for expert guidance on American and British English; and to Chris Williams for comments on Chapter 6; also to the funding agencies without whose support our research reported here would not have been possible. Thanks are due, on this score, to the Deutsche Forschungsgemeinschaft (DFG) for grant MA 1652/3 to Christian Mair and the University of Freiburg, to the Arts and Humanities Research Board (AHHR; subsequently changed to AHRC), the British Academy, and the Leverhulme Trust for research grants awarded to Geoffrey Leech at Lancaster University. We also record our gratitude to Cambridge University Press for making available to us relevant sections of the Cambridge International Corpus (CIC), and to Pearson/Longman for allowing us to consult the Longman Corpus of Spoken American English (LCSAE).
A. Abbreviations for corpora, corpus collections and subcorpora
(listed approximately in order of importance for this book)

<table>
<thead>
<tr>
<th>No.</th>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Brown</td>
<td>the Brown (University) Corpus (see 2.1, Appendix I)</td>
</tr>
<tr>
<td>2</td>
<td>LOB</td>
<td>the Lancaster–Oslo/Bergen Corpus</td>
</tr>
<tr>
<td>3</td>
<td>Frown</td>
<td>the Freiburg–Brown Corpus</td>
</tr>
<tr>
<td>4</td>
<td>F-LOB</td>
<td>the Freiburg–Lancaster–Oslo/Bergen Corpus</td>
</tr>
<tr>
<td>5</td>
<td>The Brown family</td>
<td>the four corpora above, regarded as a group</td>
</tr>
<tr>
<td>6</td>
<td>Lanc-31</td>
<td>the Lancaster 1931 Corpus, matching the four corpora above</td>
</tr>
<tr>
<td>7</td>
<td>B-LOB</td>
<td>a nickname for Lanc-31, meaning ‘before LOB’</td>
</tr>
<tr>
<td>8</td>
<td>Press,</td>
<td>Four subcorpora into which the corpora of the Brown family are divided. For the composition of the Brown corpus (and hence of the other corpora of the Brown family), see Appendix I.</td>
</tr>
<tr>
<td>9</td>
<td>General Prose,</td>
<td>the Brown family are divided. For the composition of the Brown corpus (and hence of the other corpora of the Brown family), see Appendix I.</td>
</tr>
<tr>
<td>10</td>
<td>Learned,</td>
<td>the Brown family are divided. For the composition of the Brown corpus (and hence of the other corpora of the Brown family), see Appendix I.</td>
</tr>
<tr>
<td>11</td>
<td>Fiction</td>
<td>the Brown family are divided. For the composition of the Brown corpus (and hence of the other corpora of the Brown family), see Appendix I.</td>
</tr>
<tr>
<td>12</td>
<td>BNC</td>
<td>the British National Corpus</td>
</tr>
<tr>
<td>13</td>
<td>the BNC demographic subcorpus (BNCdemog)</td>
<td>a part of the BNC, consisting of largely spontaneous spoken English discourse by 153 individuals and their interlocutors, sampled from the population of the UK on demographic principles</td>
</tr>
<tr>
<td>14</td>
<td>BNC Sampler</td>
<td>A subcorpus of the BNC, consisting of c. one million words of writing and c. one million words of speech. The POS tags are more refined than for the whole BNC, and have been post-edited for correctness.</td>
</tr>
<tr>
<td>15</td>
<td>ICE</td>
<td>the International Corpus of English</td>
</tr>
</tbody>
</table>

\(^1\) We have used the World Edition of the British National Corpus.
Abbreviations and symbolic conventions

16 ICE-GB the International Corpus of English (Great Britain) – one of the constituent corpora of ICE
17 DCPSE the Diachronic Corpus of Present-Day Spoken English
18 DSEU a mini-corpus consisting of an early part of the DCPSE
19 DICE a mini-corpus consisting of a later matching part of the DCPSE
20 ANC the American National Corpus
21 ARCHER A Representative Corpus of Historical English Registers
22 LCSAE the Longman Corpus of Spoken American English
23 CIC the Cambridge International Corpus
24 ACE the Australian Corpus of English
25 CONCE Corpus of Nineteenth-Century English
26 MICASE the Michigan Corpus of Academic Spoken English

Alphabetical index to the above list:

ACE 24 Brown family 5 Fiction 11 Lanc-31 6
ANC 20 CIC 23 F-LOB 4 LCSAE 22
ARCHER 21 CONCE 25 Frown 3 Learned 10
B-LOB 7 DCPSE 17 General Prose 9 LOB 2
BNC 12, 13, 14 DICE 19 ICE 15 MICASE 26
Brown 1 DSEU 18 ICE-GB 16 Press 8

B. Abbreviations for Geographical and Historical Subdivisions of English

AmE American English LModE Late Modern English
BrE British English ME Middle English
ModE Modern English OE Old English
EModE Early Modern English PDE Present-Day English

C. Other Abbreviations

C8 The C8 tagset: a set of part-of-speech tags used for annotating the Brown family of corpora (the C8 tags are listed in Appendix II)
CLAWS Constituent-Likelihood Automatic Word Tagging System (a POS tagger)
CLAWS4 The newest version of CLAWS
Abbreviations and symbolic conventions xxvii

CQP  Corpus Query Processor (software: a tool for interpreting corpus queries)
LL  Log likelihood (a measure of statistical significance)
NP  Noun phrase
N+N  Sequence consisting of noun + noun
N+CN  Sequence consisting of noun + common noun
OED  Oxford English Dictionary
pmw  Per million words (in statistical tables, frequencies are often normalized to this standard frequency measure)
PN+PN  A sequence of proper noun + proper noun
POS  Part of speech (used especially in the collocation ‘POS tagger/tagging’)
XML  Extensible Markup Language (an artificial metalanguage used for the encoding and processing of textual material, including corpora)

D. Conventions

[Brown L12], [LOB A09], and the like

These are address labels used to identify the whereabouts, in the Brown family of corpora, of a particular example, sentence, etc. After the corpus name, the letter indicates the text category and the two digits the number of the text sample in that category. Similar address labels are used for examples from the BNC and other corpora.

*  An asterisk before an (invented) example indicates its status as an unacceptable or ungrammatical usage.

?  A question mark before an example (invented or otherwise) indicates its questionable acceptability.

[ . . . ]  In a corpus example, an ellipsis in square brackets indicates where the example has been simplified by the omission of part of the original corpus sentence.

* ** ***  Placed next to a numerical quantity in a statistical table or bar chart, these are indicators of increasingly higher statistical significance.

* * means ‘significant at the level $p < 0.05$ (LL > 3.84)’.

** ** means ‘significant at the level $p < 0.01$ (LL > 6.63)’.

© in this web service Cambridge University Press  www.cambridge.org
xxviii Abbreviations and symbolic conventions

*** means ‘significant at the level $p < 0.001$ ($LL > 10.83$)’.

N* etc.

In referring to POS tags, an asterisk is occasionally used as a ‘wildcard symbol’, standing for any number (including zero) of characters, excepting a space or other delimiting character. For example, N* will identify any tag beginning with N, which means, in fact, any noun in the C8 tagset.

HAVE got to, SHE, HAVE, NEED to, and the like

In certain chapters, the small capitals indicate that the word cited is understood as a lemma, not as an individual word form. For example, HAVE to signifies any form of the verb HAVE followed by to (i.e. have to, has to, had to, having to). The chapters in which this convention chiefly applies are 4 and 5. It is important to avoid confusion in some contexts by using this convention. In other contexts the convention is unnecessary, as the interpretation of a graphical form like be going to is clear from the context. Hence we use this convention only in some chapters.