The Verb Phrase in English

The chapters in this volume feature new and ground-breaking research carried out by leading scholars and promising young researchers from around the world on recent changes in the English verb phrase. Drawing on authentic corpus data, the contributions consider both spoken and written English in several genres. Each contribution pays particular attention to the methodologies used for investigating short-term patterns of change in English, with detailed discussions of controversies in this area. This cutting-edge collection is essential reading for historians of the English language, syntacticians and corpus linguists.

**Bas Aarts** is Professor of English Linguistics and Director of the Survey of English Usage at University College London.

**Joanne Close** is Lecturer in English Language in the Department of English at the University of Chester.

**Geoffrey Leech** is Professor Emeritus of English Linguistics at Lancaster University.

**Sean Wallis** is a Senior Research Fellow in the Survey of English Usage at University College London.
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:

- 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*
- Geoffrey Leech, Marianne Hundt, Christian Mair and Nicholas Smith: *Change in Contemporary English: A Grammatical Study*
- Jonathan Culpeper and Merja Kytö: *Early Modern English Dialogues: Spoken Interaction as Writing*
- Daniel Schreier, Peter Trudgill, Edgar Schneider and Jeffrey Williams: *The Lesser-Known Varieties of English: An Introduction*
- Hilde Hasselgård: *Adjunct Adverbials in English*
- Raymond Hickey: *Eighteenth-Century English: Ideology and Change*
- Charles Boberg: *The English Language in Canada: Status, History and Comparative Analysis*
- Thomas Hoffmann: *Preposition Placement in English: A Usage-Based Approach*
- Claudia Claridge: *Hyperbole in English: A Corpus-Based Study of Exaggeration*
- Päivi Pahta and Andreas H. Jucker (eds.): *Communicating Early English Manuscripts*
- Irma Taavitsainen and Päivi Pahta (eds.): *Medical Writing in Early Modern English*
- Colette Moore: *Quoting Speech in Early English*
David Denison, Ricardo Bermúdez-Otero, Chris McCully and Emma Moore
(eds.): Analysing Older English
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
Earlier titles not listed are also available
The Verb Phrase in English

Investigating Recent Language Change with Corpora

Edited by

BAS AARTS
University College London

JOANNE CLOSE
University of Chester

GEOFFREY LEECH
Lancaster University

SEAN WALLIS
University College London
Contents

List of figures  ix
List of tables  xv
Notes on contributors  xx
Preface  xxvii

1 Introduction  1
BAS AARTS, JOANNE CLOSE, GEOFFREY LEECH
AND SEAN WALLIS

2 Choices over time: methodological issues in investigating current change  14
BAS AARTS, JOANNE CLOSE AND SEAN WALLIS

3 Recent shifts with three nonfinite verbal complements in English: data from the 100-million-word Time corpus (1920s–2000s)  46
MARK DAVIES

4 Verb structures in twentieth-century British English  68
NICHOLAS SMITH AND GEOFFREY LEECH

5 Nominalizing the verb phrase in academic science writing  99
DOUGLAS BIBER AND BETHANY GRAY

6 The verb phrase in contemporary Canadian English  133
SALI A. TAGLIAMONTE

7 Recent change and grammaticalization  155
MANFRED KRUG AND OLE SCHÜTZLER

8 The progressive verb in modern American English  187
MAGNUS LEVIN

9 I was just reading this article – on the expression of recentness and the English past progressive  217
MEIKE PFAFF, ALEXANDER BERGS AND THOMAS HOFFMANN
<table>
<thead>
<tr>
<th>#</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Bare infinitival complements in Present-Day English</td>
<td>239</td>
</tr>
<tr>
<td></td>
<td>Marcus Callies</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Operator and negative contraction in spoken British English</td>
<td>256</td>
</tr>
<tr>
<td></td>
<td>José Ramón Varela Pérez</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>The development of comment clauses</td>
<td>286</td>
</tr>
<tr>
<td></td>
<td>Gunther Kaltenböck</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>The perfect in spoken British English</td>
<td>318</td>
</tr>
<tr>
<td></td>
<td>Jill Bowie, Sean Wallis and Bas Aarts</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Changes in the verb phrase in legislative language in English</td>
<td>353</td>
</tr>
<tr>
<td></td>
<td>Christopher Williams</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Modals and semi-modals of obligation in American English: some aspects of developments from 1990 until the present day</td>
<td>372</td>
</tr>
<tr>
<td></td>
<td>Stig Johansson</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Notes</td>
<td>381</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>405</td>
</tr>
<tr>
<td></td>
<td>Name index</td>
<td>436</td>
</tr>
<tr>
<td></td>
<td>Subject index</td>
<td>440</td>
</tr>
</tbody>
</table>
Figures

2.1. An example tree diagram, *I think that’s fascinating* (DI-A02 #28) page 16
2.2. An FTF created with ICECUP, matching the highlighted nodes in Figure 2.1
2.3. Descending the space of possible choices to focus on true alternants
2.4. An FTF used to search for *shall* after any subject NP
2.5. An FTF for a first-person subject followed by *shall*
2.6. An FTF which specifies that *shall* may not be in VP final position
2.7a. Declining use of *shall* as a proportion $p$ of the set {shall, will} with first-person subjects, annual data, with Wilson intervals
2.7b. Declining use of *shall* as a proportion $p$ of the set {shall, will} with first-person subjects, half-decade data (‘1960’ = 1958–62 inclusive, ‘1965’ = 1963–7, etc.)
2.7c. Declining use of *shall* as a proportion $p$ of {shall, will, ’ll} with first-person subjects, annual data
2.8. An FTF for *be going to* in the same syntactic context as before
2.9. Summarising changes for *shall*, *will*, *’ll* and *be going to*, first-person positive declarative (non-VP-final) over time, plotting percentage swing ($d_p$), on an absolute (‘per million word’) and relative (‘within set’) basis, with 95 per cent confidence intervals
3.1. Decade-by-decade frequency for *[end] up* *[V-ing]* from the *Time* corpus, 1920s–2000s
3.2. Percentage of *[V-ing]* (vs *[to V]*) by verb, 1920s–2000s
3.3. Frequency of *[+for]* in the *Time* corpus, with verbs allowing alternation
4.1. Increase in negative contractions per million words (pmw): comparing the BLOB–1931, LOB, FLOB and BE06 corpora
List of figures

4.2. Increase in negative contractions as a percentage of not-negation: comparing the BLOB-1931, LOB, FLOB and BE06 corpora
4.3. Schematic diagram of the semantic field of obligation/necessity
4.4. Decline of core modals relating to obligation/necessity: frequencies pmw
4.5. Increase in frequency of have to pmw, 1931–91, in BrE, representing subcorpora as well as the overall trend
4.7. Modal + progressive combinations pmw in BrE, 1931–2006
4.8. Present progressive (active) in BrE, 1931–2006: frequencies pmw
4.9. Be-passives pmw in BrE, 1931–2006 (note: figures exclude the progressive be-passive)
4.10. Finite be-passives pmw in BrE, 1931–2006
4.11. Nonfinite be-passives pmw in BrE, 1931–2006
5.1. Historical use of nominalizations
5.2. Historical use of nominalizations in academic sub-registers
5.3. Historical use of nouns
5.4. Historical use of nouns in academic sub-registers
5.5. Historical use of nouns as nominal premodifiers
5.6. Historical use of nouns as nominal premodifiers in academic sub-registers
5.7. Historical use of simple verb phrases (no marked aspect, voice, modality) in academic registers
5.8. Historical use of perfect aspect verb phrases in academic registers
5.9. Historical use of passive voice verb phrases in academic registers
5.10. Historical use of verb phrases with modal verbs in academic registers
5.11. Historical use of relative clauses in academic registers
5.12. Historical change in the use of nouns versus verbs: academic writing
5.13. Main verbs that have decreased the most from 1900 to 2005
5.14. Historical change in the use of be-clause construction types: science research articles
6.1. Map of Toronto in the context of southeastern Canada and northeast United States
6.2. Map of Toronto showing the neighbourhoods sampled in the Toronto English Project
6.3. Toronto English Archive in comparison with the 2001 Census
List of figures xi

6.4. Proportion of have by speaker age and sex 143
6.5. Proportion of have to by speaker age and sex 143
6.6. Proportion of be going to by speaker age and sex 148
6.7. Proportion of be like by speaker age and sex 150
6.8. Composite perspective on change in progress in Toronto (both male and female subjects) 152
6.9. Generations in the Toronto English Archive 154
7.1. want to and going to per 10,000 words: a real-time study based on the British drama component of the ARCHER corpus (adapted from Krug 2000: 168) 162
7.2. Frequency (per 10,000 quotes) of going to and gonna in OED quotes (Mair 2004: 126) 162
7.3. Going to and want to in an apparent time study based on the spoken BNC (Krug 2000: 175): percentage of contracted forms 163
7.4. Idea is to/that and idea was to/that in the Time corpus (1920s–2000s): development of the index of grammaticalization composed of discourse frequency and the degree of isomorphism 169
7.5. Idea is and idea was in the Time corpus (1920s–2000s): development of the index of grammaticalization composed of discourse frequency and the degree of isomorphism, treated separately for complementation with to- and that-clauses 170
7.6. Proportion of idea is/was + to/that-clause complementation relative to all forms of idea is/was in the Time corpus (1920s to 2000s). 95 per cent confidence intervals based on Wilson-scores 170
7.7. Intention-marking constructions in the spoken BNC (c. 10.4m words): ranking according to grammaticalization index 173
7.8. Longman Pronunciation Dictionary (Wells 2008): s.v. idea 176
8.1. be being adjective in the Time corpus (pmw) 194
8.2. be being adjective in COCA and LSAC (pmw) 195
8.3. Present- and past-tense forms of be being adjective in COCA and LSAC 198
8.4. Proportion of progressives per 10,000 verbs for the set believe, (dis)like, hate, intend, know, love, pity, want and wish in the Time corpus 201
8.5. Proportion of progressives per 10,000 verbs for the set believe, (dis)like, hate, intend, know, love, pity, want and wish in COCA and LSAC 202
8.6. Progressives with expect, feel, hope, think and wonder in the Time corpus (pmw) 204
8.7. Changes in the probability of particular verbs being used with progressive aspect over time: data for expect, feel, hope, think and wonder in the Time corpus, with 'I'-shaped error bars calculated using Wilson’s score interval at a 95 per cent confidence level 204
8.8. Proportions of tenses in the Time corpus with progressive forms of expect, feel, hope, think and wonder, with 95 per cent Wilson error bars 204
8.9. Changes in the proportions of different meanings of think in the Time corpus, with 95 per cent Wilson error bars 207
8.10. Changes in the complementation patterns of progressive hope in the Time corpus (as a proportion of progressive uses), with 95 per cent Wilson error bars 209
9.1. Participant ratings for sentence (23) Funny, you should mention him, in fact I was just talking to him 212
9.2. Participant ratings for sentence (25) I was recently flying United Airlines myself, and you’re right, their leg room is really poor 226
9.3. Participant ratings for sentence (25) Someone was just telling me about this new band called Sugarplum or something 227
9.4. Participant ratings for sentence (26) Someone has just told me about this new computer – I think I might buy it 227
9.5. Participant ratings for sentence (46) I was recently flying United Airlines and you’re right, their leg room is really poor 227
9.6. Participant ratings for sentence (47) Why am I not hungry? Because I was only just having lunch 231
11.1. Overall distribution of full forms (is not), negative contraction (isn’t) and operator contraction (’s not) across time 231
11.2. Distribution of operator contraction (’s not) according to speaker age across time 269
11.3. LLC-mini-spoken. Interaction of language-internal factors influencing the choice of variant 279
11.4. ICE-GB-mini-spoken. Interaction of language-internal factors influencing the choice of variant 280
12.1. Normalised frequencies (per million words) of comment clauses I think and I suppose in LLC, ICE-GB and COLT 292
12.2. Normalised frequencies (per million words) of comment clauses I believe, I suspect, I expect, I understand, I imagine, I guess and I reckon in LLC, ICE-GB and COLT 292
12.3. Absolute pmw change of different uses of comment clause I think, including 95 per cent Newcombe-Wilson confidence intervals 293
12.4. Relative change of *I think* within the set of comment clauses overall

12.5. *I'm thinking* in the *Corpus of Historical American English* (per million words)

12.6. *I think* in the *Corpus of Historical American English* (per million words)

12.7. *I just think* in the *Corpus of Historical American English* (per million words)

12.8. *I’m guessing* in the *Corpus of Historical American English* (per million words)

12.9. Taxonomic network for comment clauses

A12.1. Per million word change over time (relative to LLC = 100%) for individual comment clauses and the total set

A12.2. Variation within the set of comment clauses: relative to the total set

A12.3. Relative change within the set of comment clauses (comparing LLC, ICE-GB, COLT), with 95 per cent Wilson score confidence intervals

A12.4. Relative change of different uses of comment clause *I think* within the set in text category B ‘informal face-to-face’ of LLC and ICE-GB, and in COLT

A12.5. Relative change of comment clause uses of *I’m thinking* compared to other uses of *I’m thinking* in the *Corpus of Historical American English*

A12.6. Relative change of comment clause uses of *I just think* and *I’m thinking* within the set (*I think*, *I just think*, *I’m thinking*) in the *Corpus of Historical American English*

13.1. Tree diagram for the sentence *I haven’t lost it*

13.2. A simple FTF to search for a present perfect auxiliary

13.3. FTF for an infinitival perfect auxiliary following a modal auxiliary under a VP

13.4. Changes in pmw frequencies (Table 13.1 ‘d%’ column) with error bars representing 95 per cent confidence intervals

13.5. FTF for perfect auxiliary as daughter of VP

13.6. FTF for perfect auxiliary preceding the VP

13.7. Changes in frequency of TPM VP categories as a proportion of all VPs

13.8. Changes in frequency of categories as a proportion of TPM VPs only

13.9. Changes in relative frequency of perfect categories against the past non-perfect

13.10. Changes in relative frequency of non-modal TPM VPs
xiv  List of figures

14.1. Frequency of *shall* per thousand words over time for each subcorpus 361
14.2. Proportion of prescriptive finite verbal constructions in the UK data for 1980 and 2010 out of the set of forms, with 95 per cent Wilson error bars 364
14.3. Proportion of prescriptive finite verbal constructions in the Australia data for 1980 and 2010 365
14.4. Distribution of actives, passives and intransitives in UK legislative texts in 1980 and 2010 367
14.5. Changing distribution of active and passive verb forms (excluding intransitive forms) in UK legislative texts from 1980 to 2010 369
15.1. The frequency of *must*, *have to*, *have got to* and *need to* (per million words, based on Smith 2003 and Leech *et al.* 2009: 283–4) 373
15.2. *must*, *have to*, *have got to* and *need to* (*Time* corpus, per million words) 374
15.3. *must*, *have to*, *have got to* and *need to* (*COCA*, per million words) 375
15.4. *must*, *have to*, *have got to* and *need to* (*COCA* text categories, per million words) 375
15.5. *I need you to* (*COCA*, per million words) 378
Tables

2.1. Change over time of ‘VP(prog)’ as a proportion of the number of words  page 18
2.2. Change over time of ‘VP(prog)’ as a proportion of the number of VPs 19
2.3. Change over time of ‘VP(prog)’ as a proportion of the number of progressivisable VPs (‘VP(+prog)’) 21
2.4. Decline in the use of shall in written corpora, LOB/F-LOB and Brown/Frown (after Mair and Leech 2006) 23
2.5a. $2 \times 2 \chi^2$ for shall and will between ICE-GB and LLC (spoken, positive and declarative; bold is significant for $p<$0.05). The contracted form ‘ll is excluded 25
2.5b. $2 \times 2 \chi^2$ for shall (‘shan’t) and will (‘ll, won’t) between LOB and FLOB (written) (data from Mair and Leech 2006) 25
2.5c. $2 \times 2 \chi^2$ for the simple lexical auxiliary verb queries for shall and will between ICE-GB and LLC, all cases, i.e. excluding the contracted form ‘ll 26
2.5d. $2 \times 2 \chi^2$ for shall and will between ICE-GB and LLC (spoken, first-person subject, declarative), excluding the contracted form ‘ll and negative cases 27
2.6a. $3 \times 2 \chi^2$ for shall/will/’ll between ICE-GB and LLC, declarative, positive, first-person subject cases ($\phi = 0.13$) 28
2.6b. A comparison of the three alternants in declarative first person, VP non-final position ($\phi = 0.11$) 29
2.7. Frequency and probability data from DCPSE reflecting a declining use of shall over time as a proportion $p(shall)$ of the set of alternants {shall, will} (left) and {shall, will, ‘ll} (right), following first-person subjects (non-VP-final) 30
2.8a. Distribution of semantic types of shall and will in first-person positive declarative utterances in DCPSE 34
2.8b. Analysis of change over time for first-person declarative Root {shall, will} 35
List of tables

2.8c. Analysis of the first-person declarative Epistemic {shall, will} alternation set over time 35

2.9. Results including the alternation for be going to in the first person 37

3.1. Phrasal verbs with up, 1970s–2000s and 1920s–50s 51

3.2. Overall frequency of V NP into V-ing, 1920s–2000s 53

3.3. Frequency of V NP into V-ing by verb, 1920s–2000s 54

3.4. New matrix verbs with V NP into V-ing construction, 1920s–2000s 55

3.5. Selected examples of V NP into V-ing, 1920s–2000s 56

3.6. to V vs V-ing by verb in the BROWN corpora 58

3.7. to V vs V-ing by verb in the Time corpus 59

3.8. [±for] with selected verbs in the Time and COCA corpora 63

3.9. Overall percentage of [±for] in the Time corpus with ask 64

3.10. Overall frequency of [+for] in the Time corpus, with verbs allowing alternation 64

4.1. The Brown family of corpora 69

4.2. Decreasing core modals and increasing emergent modals pmw in written BrE – as observed in the four Brown family corpora BLOB–1931, LOB, FLOB and BEo6 78

5.1. Overall corpus composition 103

5.2. Subcorpora within twentieth-century academic prose 104

5.3. Audience and purpose of written registers 105

5.4. Number of verb types in science research articles in 1900 and 2005 115

5.5. Verbs that have decreased in use from 1900 to 2005 115

5.6. Verbs that have increased in use from 1900 to 2005 130

6.1. Toronto neighbourhoods with the lowest proportion of immigrants 137

6.2. Quota-based sampling 138

6.3. Demographic distribution of speakers in the Toronto English Archive 139

6.4. Prescriptive rules for using will and shall 148

6.5. Changes in progress in Toronto English 152

7.1. Idea is and idea was in the Time corpus (1920s–2000s): overall frequency and proportion of to- and that- complementation vs other types of complementation 168

7.2. Frequencies of to/that-clause complementation of idea is/was in the Time corpus, comparing three selected time periods 171

7.3. Determining degrees of grammaticalization in different constructions of the type the + NP + be in the spoken BNC (c. 10.4m words) 172
7.4. Stress pattern of idea in ICE-GB
7.5. Intrusive /r/ by construction in the spoken ICE-GB
7.6. Liaison sites in the spoken BNC: idea + word beginning with vowel (orthography only)
7.7. Logistic regression predicting level or first-syllable stress
7.8. Probabilities of the occurrence of early stress under different conditions; based on Table 7.7
8.1. Progressive forms across tenses in the Time corpus with EXCEPT, FEEL, HOPE, THINK and WONDER (pmw; raw frequencies in brackets)
8.2. Meanings of progressive THINK in the Time corpus (pmw; raw frequencies in brackets)
8.3. Complementation patterns of progressive HOPE in the Time corpus (pmw; raw frequencies in brackets)
9.1. Total number of hits for recentness progressives: was/were just/recently V-ing, was/were V-ing the other Ntemp
9.2. Diachronic distribution of just V-ing recentness progressives in the COCA
9.3. Context-induced reinterpretation of the past progressive
10.1. Raw frequency counts of bare infinitival complements after selected verbs in two corpora
10.2. Distribution of bare infinitival complements in discourse domains for the three most frequent verbs (COCA)
10.3. Distribution of bare infinitival complements in discourse domains for the three most frequent verbs (BNC)
10.4. Number of bare infinitival complements according to length of intervening material between matrix and complement verb (in number of words)
10.5. Number/percentage of tokens with to preceding matrix verb
10.6. Number/percentage of tokens with neighbouring constructional trigger (i.e. another bare infinitive either immediately preceding or following the respective matrix verb)
11.1. Operator and negative contraction with is and are in ARCHER (1700–1900) (López-Cousou 2007: 307), LLC (Hiller 1987: 540) and the demographically sampled component of the BNC (Anderwald 2002: 78)
11.2. Operator and negative contraction with HAVE in ARCHER (1700–1900) (López-Cousou 2007: 312), LLC (Hiller 1987: 538) and the demographically sampled component of the BNC (Anderwald 2002: 78)
11.3. Data files used in this study
xviii List of tables

11.4. Overall distribution of full forms, negative contraction and operator contraction with have, will and would across time 266

11.5. Frequency of operator contraction for have (‘ve not/’s not/’d not), will (‘ll not) and would (‘d not) 267

11.6. Distribution of full forms, negative contraction and operator contraction with have, will and would according to age 268

11.7. Overall distribution of full forms (are not), negative contraction (aren’t) and operator contraction (‘re not) with we, you and they as subject hosts across time 268

11.8. Overall distribution of full forms (is not), negative contraction (isn’t) and operator contraction (‘s not) across time 269

11.9. Face-to-face conversation: distribution of full forms (is not), negative contraction (isn’t) and operator contraction (‘s not) across time 270

11.10. Broadcast discussions and interviews: distribution of full forms (is not), negative contraction (isn’t) and operator contraction (‘s not) across time 270

11.11. Distribution of full forms (is not), negative contraction (isn’t) and operator contraction (‘s not) according to register in the LLC sample 271

11.12. Distribution of full forms (is not), negative contraction (isn’t) and operator contraction (‘s not) according to register in the ICE-GB sample 271

11.13. Distribution of full forms (is not), negative contraction (isn’t) and operator contraction (‘s not) according to the type of subject host 273

11.14. Distribution of full forms (is not), negative contraction (isn’t) and operator contraction (‘s not) according to ellipsis of the predication 274

11.15. Distribution of full forms (is not), negative contraction (isn’t) and operator contraction (‘s not) according to the informational status of the operator in cases of ellipsis of the predication 275

11.16. Distribution of full forms (is not), negative contraction (isn’t) and operator contraction (‘s not) according to discourse function: correctives, implicit denials and explicit denials 277

11.17. Distribution of full forms (is not), negative contraction (isn’t) and operator contraction (‘s not) according to the function of the operator 278
List of tables xix

11.18. Distribution of full forms (is not), negative contraction (isn’t) and operator contraction (s not) according to speaker age 279

12.1. Frequencies of comment clauses in DCPSE (normalized per 1 million words) 290

12.2. Overall frequencies of different uses of comment clause I think in DCPSE (normalized per million words) 293

12.3. Frequencies of initial comment clauses + that (in spoken ICE-GB; from Van Bogaert 2009:378) and frequencies of verbs and their transitive use with object NP (in total ICE-GB) 310

A12.1. Different uses of comment clause I think in ICE-GB (DCPSE) text category B (informal face-to-face) and COLT (normalized per 1 million words) 315

A12.2. I’m thinking in the Corpus of Historical American English 316

13.1. Frequencies of perfect auxiliaries in DCPSE, divided by inflectional category 326

13.2. Changes in frequency of subtypes as a proportion of the set of TPM VPs, including goodness-of-fit $\chi^2$ for each subtype (‘s’ means significant at $p<0.05$) and goodness-of-fit $\phi^2$ effect size 332

13.3a. Comparing present perfect cases against present-marked non-perfect VPs 336

13.3b. Comparing present perfect cases against TPM VPs 336

13.4. Decline of past perfect usage in temporal clause contexts 341

A13.1. A worked example of Wallis’ goodness-of-fit $\phi^2$ using data from Table 13.3b 352

14.1. Frequency of shall per 1,000 words 361

14.2. Frequency of prescriptive finite verbal constructions: UK 363

14.3. Frequency of prescriptive finite verbal constructions: Australia 363

14.4. Active–passive ratios as percentages in UK 1980 texts (excluding intransitives) 368

14.5. Active–passive ratios as percentages in UK 2010 texts (excluding intransitives) 368

14.6. Most frequently used verbs in the UK texts in 1980 and in 2010 371

15.1. The top fifteen verbs following I must (COCA, sorted by ‘relevance’) 376
Notes on contributors


Alexander Bergs is Full Professor and Chair of English Language and Linguistics at the University of Osnabrück. He received his PhD from the University of Düsseldorf in 2002 where he also took his postdoctoral degree in 2006. His main areas of research include historical (socio)linguistics, language change theory, new media language, and the morphosyntax of present-day English and its varieties. He is the author of *Social Networks and Historical Sociolinguistics* (2005), *Modern Scots* (2nd edition 2005), *Synchronic English Linguistics* (2012), and editor, with Gabriele DIEWALD, of *Constructions and Language Change* (2008) and *Contexts and Constructions* (2009), and, with Laurel Brinton, of *English Historical Linguistics: An International Handbook* (2 vols., 2012).

Douglas Biber is Regents’ Professor of English (Applied Linguistics) at Northern Arizona University. His research efforts have focused on corpus linguistics, English grammar and register variation (in English and cross-linguistic; synchronic and diachronic). He has written numerous books and monographs, including *Variation across Speech and Writing* (Cambridge University Press, 1988), *Dimensions of Register Variation: A Cross-Linguistic Approach* (2004), *The Verb Phrase in English: Investigating Recent Language Change with Corpora* (2009), *dimensions of register variation: a cross-linguistic approach* (2004), and, with Mark Davies, of *The Verb Phrase in English: Investigating Recent Language Change with Corpora* (2009), and, with Laurel Brinton, of *English Historical Linguistics: An International Handbook* (2 vols., 2012).
Notes on contributors  xxi


Jill Bowie is a Research Fellow at the Survey of English Usage, UCL, where she worked on the AHRC-funded project ‘The changing verb phrase in present-day British English’. She completed a PhD on language evolution at the University of Reading in 2008. Earlier studies in linguistics were undertaken at the University of Queensland, Australia, where she also worked as a research assistant on Rodney Huddleston’s Cambridge Grammar of the English Language project. Her research interests include language evolution and change, English syntax and morphology, and the grammatical analysis of spoken discourse.

Marcus Callies studied history and English philology in Marburg, Germany and London. He holds a PhD in English Linguistics from the University of Marburg and has worked as a research assistant, junior lecturer and assistant professor at the universities of Marburg and Freiburg in Germany. He is currently Professor of English Linguistics at the University of Bremen, Germany. His research interests include grammatical variation and change in Present-Day English, (cross-cultural) variation in conceptual metaphor, and advanced learner varieties with a focus on discourse-functional, pragmatic aspects and interface relations.

Joanne Close is a Lecturer in English Language at the University of Chester. Her main research interests are in syntactic variation in English dialects and language change, particularly in contemporary English. She has worked on the AHRC-funded project ‘The changing verb phrase in present-day British English’ based at the Survey of English Usage, UCL, and her current research focuses on changes in the modal verbs and the mood system of English. She completed her PhD, ‘English auxiliaries: a syntactic account of contraction and variation’, at the University of York in 2004.

Mark Davies is a Professor of Linguistics at Brigham Young University in Provo, Utah, USA. He has published widely on historical change and genre-based variation in syntax, corpus design and use (related to English, Spanish and Portuguese), and also word frequency (with three recent frequency dictionaries). He is the creator of several large online corpora, including the 100-million-word Time Magazine Corpus of
xxii  Notes on contributors

American English (1920s–2000s), the 450-million-word Corpus of Contemporary American English (COCA, 1990–present), and the 400-million-word Corpus of Historical American English (COHA, 1810-2009) (see corpus. byu.edu).

Bethany Gray is based at Northern Arizona University. Her research efforts have focused on the application of corpus research methodologies to study language use in academic writing, with articles published in the Journal of English for Academic Purposes and several edited book chapters.

Thomas Hoffmann is Assistant Professor at the University of Osnabrück. His main research interests are usage-based Construction Grammar, synchronic syntactic variation and World Englishes. He has published articles in international journals such as Cognition Linguistics, English World-Wide and Corpus Linguistics and Linguistic Theory, and his monograph Preposition Placement in English was published by Cambridge University Press in 2011. With Graeme Trousdale he edited the Oxford Handbook of Construction Grammar (2013), and with Alexander Bergs he is an editor of the online journal Constructions.

Stig Johansson was Professor of Modern English Language at the University of Oslo from 1976 until his retirement in 2008. His contributions to corpus linguistics are many and varied, but perhaps it will be for his initiative in co-founding the International Computer Archive in Modern and Medieval English (ICAME), of which he was the first Coordinating Secretary (1977–96), that he will be most remembered. Among his many achievements, Stig was a pioneer in co-developing the Lancaster–Oslø/Bergen Corpus (LOB), the first computer corpus of British English. He went on to take a leading role in developing the English–Norwegian Parallel Corpus (ENPC), the Oslo Multilingual Corpus (OMC) and the Norwegian component of the International Corpus of Learner English (NICLE). As a great corpus grammarian he was meticulous and insightful, as is seen particularly in his major contribution to the influential Longman Grammar of Spoken and Written English (by Douglas Biber et al. 1999). In the last decade of his life, his interests increasingly turned to contrastive linguistics, through his creation and investigation of parallel and multilingual corpora.

Gunther Kaltenböck studied English and French philology in Vienna and London and holds an MA in Modern English Language from the University of London and a PhD in English linguistics from the University of Vienna. He currently works as Senior Lecturer in English Linguistics at the University of Vienna. His research interests include cognitive-functional grammar, grammatical variation, phonetics and pragmatics.

Manfred Krug is Chair of English and Historical Linguistics at the University of Bamberg, Germany. In his research, he focuses on
frequency-driven approaches to language variation and change, typically from a grammaticalization perspective. He holds an MA in Applied Linguistics from Exeter University and studied Latin, English and German at Freiburg University, where he completed his PhD and postdoctoral research project (Habilitation). He was a visiting professor at Portland State University (Oregon, USA) and has held the chair of English Historical Linguistics at the University of Mannheim.

Geoffrey Leech is Emeritus Professor of English Linguistics at Lancaster University (England), where he has been a faculty member for forty years. He has published many books and articles in the fields of English grammar, stylistics, pragmatics, semantics and corpus linguistics. His research career dates from the 1960s, when he produced three books (English in Advertising, A Linguistic Guide to English Poetry and Towards a Semantic Description of English), and extends to the present time: his recent works include Language in Literature (2008) and Change in Contemporary English: A Grammatical Study (with Marianne Hundt, Christian Mair and Nicholas Smith, 2009). Among his best-known books are Principles of Pragmatics (1983) and A Comprehensive Grammar of the English Language (with Randolph Quirk, Sidney Greenbaum and Jan Svartvik, 1985). In the 1970s, he pioneered the development of computer corpus linguistics, and he led the Lancaster team in the compilation of the British National Corpus (BNC), which has been a key tool in the advancement of English Linguistics for the past fifteen years. He studied at MIT for a year in 1964–5, and has held visiting positions in Australia, China, France, Japan, New Zealand and the USA. He has also visited over thirty countries as an invited lecturer.

Magnus Levin received his PhD from Lund University in 2001 and has since then been working as a Senior Lecturer in Växjö. His research interests mainly include the study of change and variation in different varieties of English and the exploration of phraseology in Present-Day English.

José Ramón Varela Pérez is based at the University of Santiago de Compostela. His doctoral dissertation is on variation and change in negative constructions in contemporary British English. His research interests include the study of current grammatical and lexical change in English, historical linguistics, TEFL and SLA.

Meike Pfaff is a Lecturer in English Linguistics at the University of Hannover. She previously worked at the University of Osnabrück where she also received her PhD. Her doctoral thesis was on selected phenomena of tense, aspect and modality from a Construction Grammar point of view. Other research interests of hers include language change and language contact as well as varieties of English. Her master’s thesis, which she wrote under the supervision of Tania Kuteva at the University of Düsseldorf, was
xxiv  Notes on contributors

a comparative case study of contact-induced language change, investigating
the different relativization strategies of Irish English and Scottish English.

Ole Schützler  studied English Linguistics and Literature at the
University of Mannheim as well as Music at the University of Music
and Performing Arts, Mannheim. He currently works as an assistant profes-
sor in the Department of English Linguistics at the University of Bamberg,
where he has recently completed his PhD on phonological and phonetic
variation in Scottish English. His other research interests comprise varieties
(and variation) of English in general (especially under the aspect of socio-
phonetics/accent studies), acoustic phonetics and methods of statistical data
analysis.

Nicholas Smith  is a Lecturer in English Language and Linguistics at the
University of Salford. His research and teaching interests are in the use of
computer corpora to study variation and change in the English language. In
the 1990s, while at Lancaster University, he worked on enhancements to the
software used for the grammatical tagging of the British National Corpus. He has
published numerous articles on methods of corpus annotation and analysis, and
pioneered the compilation of new corpora, including two early twentieth-century
corpora of British English that use the sampling frame of the 1960s Brown and
LOB corpora. He has used these corpora extensively for the study of current
change in English, and published the results in refereed journal articles and most
recently in Change in Contemporary English: A Grammatical Study (with Geoffrey
Leech, Christian Mair and Marianne Hundt, 2009). He also has considerable
experience of teaching and tutoring in corpus linguistics. He has co-authored an
introductory guide in this field, Corpus Linguistics with BNCweb – a Practical
Guide (2008), which was shortlisted for the BAAL book prize.

Sali A. Tagliamonte  is a Professor in the Linguistics Department at the
University of Toronto, Canada. She is the author of Analysing Sociolinguistic
Variation (Cambridge University Press, 2006) and co-author of African
American English in the Diaspora: Tense and Aspect (2001). She has published
on African American varieties, British, Irish and Canadian dialects, teen
language and television. Her ongoing research focuses on morphosyntactic
discourse-pragmatic features using cross-variety and apparent-time
comparisons in synchronic corpora to explore linguistic change.

Sean Wallis  is a Senior Research Fellow in the Survey of English Usage at
University College London. He specializes in the development of advanced
research methods in corpus linguistics, with a focus on parsed corpora. He is
the author of the International Corpus of English Corpus Utility Program
(ICECUP 3.1 and iv) and worked on the compilation of the British com-
ponent of the International Corpus of English (ICE-GB) and the Diachronic
Corpus of Present-Day Spoken English (DCPSE). He has written a number
of articles on corpus linguistics methodology, is the editor of the statistics
Notes on contributors  xxv

Christopher Williams  is Full Professor of English at the University of Foggia, Italy. His main research interests are in the field of tense, aspect and modality, as well as in legal English. His monographic works include Non-progressive and Progressive Aspect in English (2002) and Change and Tradition in Legal English: Verbal Constructions in Prescriptive Texts (2005). He co-edited (with Ilse Depraetere) a volume on future time reference in English, which appeared as a special issue of English Language and Linguistics (vol. 14 (2), July 2010). He is also editor of the journal ESP Across Cultures.
The origin of this book can be traced back to a symposium held at the University of London on Tuesday, 14 July 2009, organized by Bas Aarts and Joanne Close in celebration of the fiftieth anniversary of the founding of the Survey of English Usage (SEU) by Randolph Quirk (now Lord Quirk) at University College London. We were fortunate that Randolph himself was able to attend part of the symposium, and to listen to an address entitled ‘Surviving surveying’ by David Crystal, who light-heartedly reminisced on the early days of the Survey in the 1960s when spoken materials which eventually became part of the SEU Corpus, the London Lund Corpus (LLC), the ICE-GB Corpus and the Diachronic Corpus of Present-Day Spoken English (DCPSE) were transcribed.

The majority of papers given at the symposium have been reworked as contributions to this edited collection. As often happens with symposia, not all the major figures active in this research field could attend, but we were subsequently able to invite contributions from other international leaders in the field such as Douglas Biber and Mark Davies. These have enhanced the present volume’s representation of the state of current research in this fast-developing field.

The editors would like to thank Merja Kytö, General Editor of the Studies in English Language series, Helen Barton at Cambridge University Press, our copy-editor Kay McKechnie and proofreader Michael Sheppard for their help along the way.

We are also extremely grateful to a number of colleagues who acted as readers of draft versions of the chapters in this book, namely Lieselotte Anderwald, David Denison, Johan Elsness, Sebastian Hoffmann, Marianne Hundt, Alison Johnson, Svenja Kranich, Christian Mair, Paloma Núñez-Pertejo, Jennifer Smith, Erik Smitterberg and Anthony Warner.