CORPUS LINGUISTICS
Critical Concepts in Linguistics  
Series editor: Carlos P. Otero  
Other titles in this series:

Morphology  
Edited with a new introduction by Francis Katamba  
6 volume set

Lexicography  
Edited with a new introduction by Reinhard Hartmann

Semantics  
Edited with a new introduction by Javier Gutiérrez-Rexach  
6 volume set

Pragmatics  
Edited with a new introduction by Asa Kasher  
6 volume set

Phonology  
Edited with a new introduction by Charles W. Kreidler  
6 volume set

Critical Discourse Analysis  
Edited with a new introduction by Michael Toolan  
4 volume set

World Englishes  
Edited with a new introduction by Kingsley Bolton and  
Braj B. Kachru  
6 volume set

Syntax  
Edited with a new introduction by Robert Freidin and  
Howard Lasnik  
6 volume set

Forthcoming:

Lexicology  
Edited with a new introduction by Patrick Hanks  
6 volume set

Sociolinguistics  
Edited with a new introduction by Nikolas Coupland and  
Adam Jaworski  
6 volume set

Language Acquisition  
Edited with a new introduction by Charles Yang  
4 volume set
CORPUS LINGUISTICS

Critical Concepts in Linguistics

Edited by
Wolfgang Teubert
&
Ramesh Krishnamurthy

Volume I

Routledge
Taylor & Francis Group
LONDON AND NEW YORK
Dedication

In fond memory of John Sinclair (1933–2007) a pioneer of corpus linguistics and an illuminating mentor, colleague, and friend
CONTENTS

Acknowledgements xix
Chronological table of reprinted articles and chapters xxiii
Foreword xxxvii

General introduction 1

VOLUME I

PART 1

Theoretical aspects of corpus linguistics 39

1 The importance of corpus linguistics to understanding the nature of language 41
   WALLACE CHAFE

2 Does corpus linguistics exist? Some old and new issues 58
   JAN AARTS

3 The corpus-driven approach 74
   ELENA TOGNINI-BONELLI

4 “Corporate bridges” twixt text and language 93
   ROBERT DE BEAUGRANDE

5 Corpora, databases and the organisation of linguistic data 119
   GERRY KNOWLES

6 Writing, hermeneutics, and corpus linguistics 134
   WOLFGANG TEUBERT

7 Apples and oranges: on comparing data from different corpora 160
   HANS LINDQUIST AND MAGNUS LEVIN
CONTENTS

8 Corpus linguistics, Chomsky and fuzzy tree fragments 173
   BAS AARTS

9 Meaning in the framework of corpus linguistics 182
   JOHN SINCLAIR

10 “Corpus linguistics” or “computer-aided armchair linguistics” 197
   CHARLES FILLMORE

PART 2
History of corpus linguistics 221

11 English lexical collocations: a study in computational
   linguistics 223
   JOHN SINCLAIR AND SUSAN JONES

12 Interview with John Sinclair conducted by Wolfgang
   Teubert 270
   JOHN SINCLAIR AND WOLFGANG TEUBERT

13 Problems of assembling and computerizing large corpora 285
   W. NELSON FRANCIS

14 Towards a description of English usage 299
   RANDOLPH QUIRK

15 The value of a corpus in English language research: a reappraisal 315
   GEOFFREY LEECH

16 Claimed and unclaimed sources of corpus linguistics 326
   JACQUELINE LÉON

VOLUME II

Acknowledgements ix

PART 3
Corpus composition and compilation 1

17 Corpora 3
   GEOFFREY LEECH
CONTENTS

18 Compiling and using the IJS-ELAN parallel corpus 18
   TOMAŽ ERJAVEC

19 Where did we go wrong? A retrospective look at the
   British National Corpus 35
   LOU BURNARD

20 Towards an English–Norwegian parallel corpus 55
   STIG JOHANSSON AND KNUT HOFLAND

21 The American National Corpus: overall goals and the
   first release 69
   RANDI REPPEN AND NANCY IDE

22 Parallel corpora: a real-time approach to the study of
   language change in progress 77
   CHRISTIAN MAIR

23 The development of the International Corpus of English
   91
   SIDNEY GREENBAUM

24 Corpus design criteria 99
   SUE ATKINS, JEREMY CLEAR AND NICHOLAS OSTLER

25 Representativeness in corpus design 134
   DOUGLAS BIBER

26 The computer learner corpus: a versatile new source of data
   for SLA research 166
   SYLVIANE GRANGER

27 Evolution and present situation of corpus research in China
   183
   ZHIWEI FENG

PART 4
Standardisation, alignment, tagging and corpus related
software 217

28 The detection of inconsistency in manually tagged text
   219
   HANS VAN HALTEN

29 Comparing corpora 232
   ADAM KILGARRIFF

30 From the Rosetta stone to the information society: a survey
   of parallel text processing 264
   JEAN VÉRONIS
CONTENTS

31 Char_align: a program for aligning parallel texts at the character level
KEN CHURCH
290
32 PC analysis of key words – and key key words
MIKE SCOTT
303
33 AUTASYS: grammatical tagging and cross-tagset mapping
ALEX C. FANG
318
34 Meaning and interpretation of markup
MICHAEL SPERBERG-McQUEEN, CLAUS HUITFELDT AND ALLEN RENEAR
334
35 A formal framework for linguistic annotation
STEVEN BIRD AND MARK LIBERMAN
355
36 A framework of a mechanical translation between Japanese and English by analogy principle
MAKOTO NAGAO
407
37 Intuition and annotation: the discussion continues
JOHN SINCLAIR
415
38 From Firth principles: computational tools for the study of collocation
JEREMY CLEAR
436
39 Bi- and trilingual alignment and concordancing as machine aids to human translation
MICHAEL P. OAKES AND BARBARA LEWANDOWSKA-TOMASZCZYK
457
40 One sense per discourse
WILLIAM GALE, KENNETH CHURCH AND DAVID YAROWSKY
467
CONTENTS

VOLUME III

Acknowledgements vii

PART 5

Lexicography, collocation, idioms and phraseology 1

41 The search for units of meaning 3
   JOHN SINCLAIR

42 Irony in the text or insincerity in the writer?
   The diagnostic potential of semantic prosodies 30
   BILL LOUW

43 Semantic prosodies in English and Portuguese:
   A contrastive study 51
   TONY BERBER SARDINHA

44 Metaphorical polysemy and paradigmatic relations:
   a corpus study 67
   ALICE DEIGNAN

45 Patterns of word usage in corpus linguistics 85
   RŪTA MARCINKΕVΙČIΕNĖ

46 The weight of words: an investigation of lexical gravity 97
   OLIVER MASON

47 Corpus-based lexicon building: an overview across projects,
   problems, approaches 112
   NICOLETTA CALZOLARI

48 Enthusiasm and condescension: changing norms of
   meaning and use 140
   PATRICK HANKS

49 The distribution of idioms in English 154
   ROSAMUND MOON

50 Collocations and semantic profiles: on the cause of
   the trouble with quantitative studies 166
   MICHAEL STUBBS

51 Asian or Western realities? Collocations in
   Singaporean-Malaysian English 194
   VINCENT OOI
CONTENTS

52 Lexis and lexicographers: the vocabulary of dictionary definitions
GEOFF BARNBROOK

53 On the phraseology of spoken English: the evidence of recurrent word-combinations
BENGT ALLENBERG

54 Textual colligation: a special kind of lexical priming
MICHAEL HOEY

55 Lexicographic relevance: selecting information from corpus evidence
SUE ATKINS, CHARLES FILLMORE AND CHRISTOPHER JOHNSON

PART 6
Terminology

56 Towards a corpus-based approach to terminography
LYNNE BOWKER

57 In search of representativity in specialised corpora: categorisation through collocation
GEOFFREY WILLIAMS

58 Making a workable glossary out of a specialised corpus: term extraction and expert knowledge
CHRISTINE CHODKIEWICZ, DIDIER BOURIGAULT AND JOHN HUMBLEY

59 Comment accéder aux éléments définitoires dans les textes spécialisés?
JENNIFER PEARSON
CONTENTS

VOLUME IV

Acknowledgements ix

PART 7
Grammar 1

60 Looking at looking: functions and contexts of progressives in spoken English and ‘school’ English 3
UTE RÖMER

61 The textlinguistic dimension of corpus linguistics: the support function of English general nouns and its theoretical implications 15
MICHAELA MAHLBERG

62 Between and through: the company they keep and the functions they serve 26
GRAEME KENNEDY

63 What is a grammatical rule? 41
DIETER MINDT

64 Will corpus linguistics revolutionize grammar teaching in the 21st century? 55
SUSAN CONRAD

65 Why grammar is beyond belief 67
MICHAEL HOEY

66 A corpus-driven approach to grammar: principles, methods and examples 84
GILL FRANCIS

67 A quantitative study of polarity and primary tense in the English finite clause 104
MICHAEL A. K. HALLIDAY AND ZOE JAMES

68 Subordinate clauses in English 138
JOHN KIRK

69 Lexical gaps 154
GÖRAN KJELLNER
PART 8
Translation studies, multilingual and parallel corpora 165

70 Using parallel texts in the translator training environment 167
   JENNIFER PEARSON

71 Lost in parallel concordances 176
   ANA FRANKENBERG-GARCIA

72 Thematic structure in translation between English and Norwegian 191
   HILDE HASSELGÅRD

73 Contrastive linguistics and corpora 213
   STIG JOHANSSON

74 Adverbial connectors in English and Swedish: semantic and lexical correspondences 227
   BENGT ALTENBERG

75 Epistemic possibility in an English-Swedish contrastive perspective 248
   KARIN AIJMER

76 Translators at play: exploitations of collocational norms in German–English translation 271
   DOROTHY KENNY

77 Parallel corpora in translation studies: issues in corpus design and analysis 285
   FEDERICO ZANETTIN

78 Fishing for translation equivalents using grammatical anchors 299
   TAMÁS VÁRADI

79 Une nouvelle approche à l’extraction de lexiques bilingues à partir de corpus comparables 313
   HERVÉ DÉJEAN AND ÉRIC GAUSSIER

80 Naturalness and contrastive linguistics 336
   RAPHAEL SALKIE

81 Fingerprints in translation 352
   MARTIN GELLERSTAM
CONTENTS

82 A preliminary analysis of collocational differences in monolingual comparable corpora 366
MARCO BARONI AND SILVIA BERNARDINI

83 A corpus-based view of similarity and difference in translation 384
MONA BAKER

VOLUME V

Acknowledgements vii

PART 9
Critical discourse analysis / evaluation / stylistics / rhetoric 1

84 Speaking and writing in the university: a multidimensional comparison 3
DOUGLAS BIBER, SUSAN CONRAD, RANDI REPPEN,
PAT BYRD AND MARIE HELT

85 Description and interpretation in critical discourse analysis 42
JOHN FLOWERDEW

86 Corpus evidence on Australian style and usage 54
PAM PETERS

87 Corpus, comparison, culture: doing the same things differently in different cultures 68
GEOFF THOMPSON

88 Footing shift for attribution: ‘According to the New York Times this morning . . .’ 88
ALAN PARTINGTON

89 Modal verbs in academic writing 109
PAUL THOMPSON

90 Compressed noun-phrase structures in newspaper discourse: the competing demands of popularization vs. economy 130
DOUGLAS BIBER

91 A province of a federal superstate ruled by an unelected bureaucracy: keywords of the Euro-sceptic discourse in Britain 142
WOLFGANG TEUBERT
CONTENTS

92 Ethnic, racial and tribal: the language of racism? 179
RAMESHI KRISHNAMURTHY

93 On the use of corpora in the analysis of forensic texts 201
MALCOLM COULTHARD

94 Domestic discord, rocky relationships: semantic prosodies in representations of marital violence in the O.J. Simpson trial 219
JANET COTTERILL

95 Corpus-based analysis of evaluative lexis 244
JOANNA CHANNELL

PART 10
Language history / historical linguistics 265

96 Gender differences in the evolution of standard English: evidence from the Corpus of Early English Correspondence 267
TERTTU NEVALAINEN

97 Politeness and modal meaning in the construction of humiliative discourse in an early eighteenth-century network of patron–client relationships 288
SUSAN FITZMAURICE

98 Grammaticalisation from side to side: on the development of beside(s) 318
MATTI RISSANEN

99 Advanced research on syntactic and semantic change with the Corpus del Español 337
MARK DAVIES

100 Collocational and idiomatic aspects of verbs in Early Modern English: a corpus-based study of MAKE, HAVE, GIVE, TAKE, and DO 349
MERJA KYTÖ

101 Three changing patterns of verb complementation in Late Modern English: a real-time study based on matching text corpora 386
CHRISTIAN MAIR
CONTENTS

VOLUME VI

Acknowledgements vii

PART II
Language teaching 1

102 Should you be persuaded: two samples of data-driven learning materials 3
TIM JOHNS

103 The pedagogical value of native and learner corpora in EFL grammar teaching 23
FANNY MEUNIER

104 A bird’s-eye view of learner corpus research 44
SYLVIANE GRANGER

105 Cross-disciplinary comparisons of hedging: some findings from the Michigan Corpus of Academic Spoken English 73
DEANNA POOS AND RITA SIMPSON

106 English writing by Dutch–speaking students 92
PIETER DE HAAN

107 Corpora in the classroom: an overview and some reflections on future developments 102
SILVIA BERNARDINI

108 A new academic word list 123
AVERIL COXHEAD

109 A corpus-based approach to Modern Greek language research and teaching 150
DIONYSIS GOUTSOS, OURANIA HATZIDAKI AND PHILIP KING

110 Norms for the Indian English classroom: a corpus-linguistic perspective 157
JOYBRATO MUKHERJEE

xvii
## CONTENTS

**PART 12**

**Spoken language / discourse studies**

111 Integrated and fragmented worlds: EAP materials and corpus linguistics  
JOHN SWALES  
177

112 Engaged listenership in spoken academic discourse: the case of student–tutor meetings  
FIONA FARR  
196

113 The corpus of English as lingua franca in academic settings  
ANNA MAURANEN  
217

114 The performance of speech acts in workplace conversations and the teaching of communicative functions  
ALMUT KOESTER  
232

115 Common language: corpus, creativity and cognition  
RONALD CARTER  
251

116 Functional grammar and discourse studies  
CHARLES MEYER  
276

117 Talking back: “Small” interactional response tokens in everyday conversation  
MICHAEL McCARTHY  
292

118 Spoken corpus design  
STEVE CROWDY  
320

119 The spoken language corpus: a foundation for grammatical theory  
MICHAEL A. K. HALLIDAY  
334

*Index*  
363
ACKNOWLEDGEMENTS

We thank Routledge for providing us with this unique opportunity, and the authors and publishers of the articles in this collection for their permission to reprint them. We gratefully acknowledge the assistance of several eminent scholars in the field who commented on our initial contents list: John Sinclair, Douglas Biber, Graeme Kennedy, Michael Stubbs, Bengt Altenberg, and Patrick Hanks. Unfortunately, for various logistical reasons, we were not always able to implement all of their suggestions. The final selection is therefore the sole responsibility of the co-editors, and we apologise heartily to the numerous authors whose significant articles we were unable to include in the final version.

We would like to thank Simon Alexander (of Routledge) for the support we received from the publisher. And finally, we express our sincerest gratitude to María Oset García in the early stages, and Iztok Kosem during the final stages, for their administrative assistance.

The publishers would like to thank the following for permission to reprint their material:


Robert de Beaugrande for permission to reprint Robert de Beaugrande, ‘“Corporate Bridges” Twixt Text and Language’, www.beaugrande.com/
ACKNOWLEDGEMENTS

Corporate%20Bridges.htm. specially edited version for this publication, pp. 1–18.


xx
ACKNOWLEDGEMENTS


Disclaimer

The publishers have made every effort to contact authors/copyright holders of works reprinted in the Corpus Linguistics: Critical Concepts in Linguistics. This has not been possible in every case, however, and we would welcome correspondence from those individuals/companies whom we have been unable to trace.
### Chronological table of reprinted articles and chapters

<table>
<thead>
<tr>
<th>Date</th>
<th>Author</th>
<th>Title</th>
<th>Source</th>
<th>Vol.</th>
<th>Chap.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1982</td>
<td>W. Nelson Francis</td>
<td>Problems of assembling and computerizing large corpora</td>
<td></td>
<td>I</td>
<td>13</td>
</tr>
<tr>
<td>Date</td>
<td>Author</td>
<td>Title</td>
<td>Source</td>
<td>Vol.</td>
<td>Chap.</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------</td>
<td>-----------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>1992</td>
<td>Sue Atkins, Jeremy Clear and Nicholas Ostler</td>
<td>Corpus design criteria</td>
<td><em>Literary and Linguistic Computing</em> 7:1, 1–16</td>
<td>II</td>
<td>24</td>
</tr>
<tr>
<td>1992</td>
<td>Wallace Chafe</td>
<td>The importance of corpus linguistics to understanding the nature of language</td>
<td>J. Svartvik (ed.), <em>Directions in Corpus Linguistics</em>: Proceedings of Nobel Symposium 82, Berlin: Mouton de Gruyter, pp. 79–97</td>
<td>I</td>
<td>1</td>
</tr>
<tr>
<td>1992</td>
<td>William Gale, Kenneth Church and David Yarowsky</td>
<td>One sense per discourse</td>
<td><em>Spoken corpus design: a program for aligning parallel texts at the character level</em></td>
<td>II</td>
<td>40</td>
</tr>
<tr>
<td>1993</td>
<td>Ken Church</td>
<td>Char_align: a program for aligning parallel texts at the character level</td>
<td><em>Proceedings of the 31st Annual Conference of the Association for Computational Linguistics</em>, Columbus: Association for Computational Linguistics, pp. 1–8</td>
<td>II</td>
<td>31</td>
</tr>
<tr>
<td>1993</td>
<td>Jeremy Clear</td>
<td>From Firth principles: computational tools for the study of collocation</td>
<td><em>Literary and Linguistic Computing</em> 8:4, 243–257</td>
<td>II</td>
<td>38</td>
</tr>
<tr>
<td>1993</td>
<td>Steve Crowdy</td>
<td>Spoken corpus design</td>
<td><em>Spoken corpus design: A quantitative study of polarity and primary tense in the English finite clause</em></td>
<td>VI</td>
<td>118</td>
</tr>
<tr>
<td>1993</td>
<td>Michael A. K. Halliday and Zoe James</td>
<td>Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies</td>
<td><em>Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies</em></td>
<td>III</td>
<td>42</td>
</tr>
</tbody>
</table>
1995  Michael Stubbs  Collocations and semantic profiles: on the cause of the trouble with quantitative studies  Functions of Language 2:1, 23–55  III  50
1996  Lynne Bowker  Towards a corpus-based approach to terminography  Terminology 3:1, 27–52  III  56
1997  Christian Mair  Parallel corpora: a real-time approach to the study of language change in progress  Corpus-Based Studies in English, Amsterdam: Rodopi, pp. 195–209  II  22
<table>
<thead>
<tr>
<th>Date</th>
<th>Author</th>
<th>Title</th>
<th>Source</th>
<th>Vol</th>
<th>Chap</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>Mike Scott</td>
<td>PC analysis of key words – and key words</td>
<td><em>System</em> 25:2, 233–245</td>
<td>II</td>
<td>32</td>
</tr>
<tr>
<td>1999</td>
<td>Alice Deignan</td>
<td>Metaphorical polysemy and paradigmatic relations: a corpus study</td>
<td><em>Word</em> 50:3, 319–338</td>
<td>III</td>
<td>44</td>
</tr>
<tr>
<td>Year</td>
<td>Author(s)</td>
<td>Title</td>
<td>Journal/Book/Conference Details</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>John Flowerdew</td>
<td>Description and interpretation in critical discourse analysis</td>
<td><em>Journal of Pragmatics</em> 31:8, 1089–1099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>Merja Kytö</td>
<td>Collocational and idiomatic aspects of verbs in Early Modern English: a corpus-based study of MAKE, HAVE, GIVE, TAKE, and DO</td>
<td><em>Terminologies nouvelles</em> 19, 21–28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>Bas Aarts</td>
<td>Corpus linguistics, Chomsky and fuzzy tree fragments</td>
<td><em>Corpus Linguistics and Linguistic Theory</em>, Amsterdam: Rodopi, pp. 5–13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>Averil Coxhead</td>
<td>A new academic word list</td>
<td><em>TESOL Quarterly</em> 34:2, 213–238</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>Hans Lindquist and Magnus Levin</td>
<td>Apples and oranges: on comparing data from different corpora</td>
<td><em>Corpus Linguistics and Linguistic Theory</em>, Amsterdam: Rodopi, pp. 201–214</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Author</td>
<td>Title</td>
<td>Source</td>
<td>Vol.</td>
<td>Chap.</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>2000</td>
<td>Rūta Marcinkevičienė</td>
<td>Patterns of word usage in corpus linguistics</td>
<td>Kalbotyra 49:3, 71–80</td>
<td>III</td>
<td>45</td>
</tr>
<tr>
<td>2000</td>
<td>Tony Berber Sardinha</td>
<td>Semantic prosodies in English and Portuguese: a contrastive study</td>
<td>Cuadernos de Filología Inglesa (Spain: Murcia) 9:1, 93–110</td>
<td>III</td>
<td>43</td>
</tr>
<tr>
<td>2000</td>
<td>Michael Sperberg-McQueen, Claus Huitfeldt and Allen Renear</td>
<td>Meaning and interpretation of markup</td>
<td>Markup Languages: Theory and Practice 2:3, 215–234</td>
<td>II</td>
<td>34</td>
</tr>
<tr>
<td>2000</td>
<td>Tamás Váradi</td>
<td>Fishing for translation equivalents using grammatical anchors</td>
<td>International Journal of Corpus Linguistics 5:1, 1–16</td>
<td>IV</td>
<td>78</td>
</tr>
<tr>
<td>Year</td>
<td>Author(s)</td>
<td>Title</td>
<td>Journal/Publication Details</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>Steven Bird and Mark Liberman</td>
<td>A formal framework for linguistic annotation</td>
<td><em>Speech Communication</em> 33:1–2, 23–60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>Rosamund Moon</td>
<td>The distribution of idioms in English</td>
<td>S. Nuccorini (ed.), <em>When a Torch Becomes a Candle: Variation in Phraseology</em>, SILTA, 30:2, 229–241</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>Geoff Thompson</td>
<td>Corpus, comparison, culture: doing the same things differently in different cultures</td>
<td>M. Ghadessy, R. Roseberry and A. Henry (eds.), <em>Small Corpus Studies and ELT: Theory and Practice</em>, Amsterdam: John Benjamins, pp. 311–334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>Elena Tognini-Bonelli</td>
<td>The corpus-driven approach</td>
<td>Elena Tognini-Bonelli, <em>Corpus Linguistics at Work</em>, Amsterdam: John Benjamins, pp. 84–100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Jan Aarts</td>
<td>Does corpus linguistics exist? Some old and new issues</td>
<td>L. E. Breivik and A. Hasselgren (eds.), <em>From the COLT’s mouth… and others: Language corpora studies in honour of Anna-Brita Stenström</em>, Amsterdam: Rodopi, pp. 1–16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Douglas Biber, Susan Conrad, Randi Reppen, Pat Byrd and Marie Helt</td>
<td>Speaking and writing in the university: a multidimensional comparison</td>
<td><em>TESOL Quarterly</em> 36:1, 9–48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Author</td>
<td>Title</td>
<td>Source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Tomáš Erjavec</td>
<td>Compiling and using the IJS-ELAN parallel corpus</td>
<td><em>Informatica</em> 26:3, 299–307</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Susan Fitzmaurice</td>
<td>Politeness and modal meaning in the construction of humiliative discourse in an early eighteenth-century network of patron–client relationships</td>
<td><em>English Language and Linguistics</em> 6:2, 239–265</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Christian Mair</td>
<td>Three changing patterns of verb complementation in Late Modern English: a real-time study based on matching text corpora</td>
<td><em>English Language and Linguistics</em> 6:1, 105–131</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>Author(s)</td>
<td>Title</td>
<td>Page</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Fanny Meunier</td>
<td>The pedagogical value of native and learner corpora in EFL grammar teaching</td>
<td>103</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Charles Meyer</td>
<td>Functional grammar and discourse studies</td>
<td>116</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Dieter Mindt</td>
<td>What is a grammatical rule?</td>
<td>63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Joybrato Mukherjee</td>
<td>Norms for the Indian English classroom: a corpus-linguistic perspective</td>
<td>110</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Michael P. Oakes and Barbara Lewandowska-Tomaszczyk</td>
<td>Bi- and trilingual alignment and concordancing as machine aids to human translation</td>
<td>39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Deanna Poos and Rita Simpson</td>
<td>Cross-disciplinary comparisons of hedging: some findings from the Michigan Corpus of Academic Spoken English</td>
<td>105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Paul Thompson</td>
<td>Modal verbs in academic writing</td>
<td>89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>Geoffrey Williams</td>
<td>In search of representativity in specialised corpora: categorisation through collocation</td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>Sue Atkins, Charles Fillmore and Christopher Johnson</td>
<td>Lexicographic relevance: selecting information from corpus evidence</td>
<td>55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chronological Table continued

<table>
<thead>
<tr>
<th>Date</th>
<th>Author</th>
<th>Title</th>
<th>Source</th>
<th>Vol.</th>
<th>Chap.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>Ana Frankenberg-Garcia</td>
<td>Lost in parallel concordances</td>
<td>G. Aston, S. Bernardini and D. Stewart (eds.), <em>Corpora in Translation Education</em>, Manchester: St Jerome, pp. 15–24</td>
<td>IV</td>
<td>71</td>
</tr>
<tr>
<td>2003</td>
<td>Michael Hoey</td>
<td>Why grammar is beyond belief</td>
<td><em>Belgian Journal of English Language and Literatures</em>, Special Issue: 153–196</td>
<td>IV</td>
<td>65</td>
</tr>
<tr>
<td>2003</td>
<td>Stig Johansson</td>
<td>Contrastive linguistics and corpora</td>
<td>S. Granger, J. Lerot and S. Petch-Tyson (eds.), <em>Corpus-Based Approaches to Contrastive Linguistics and Translation Studies</em>, Amsterdam: Rodopi, pp. 31–44</td>
<td>IV</td>
<td>73</td>
</tr>
<tr>
<td>Year</td>
<td>Author</td>
<td>Title</td>
<td>Journal/Publication Details</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>Göran Kjellmer</td>
<td>Lexical gaps</td>
<td>System 30:2, 167–184</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>Almut Koester</td>
<td>The performance of speech acts in workplace conversations and the teaching of communicative functions</td>
<td>International Journal of Corpus Linguistics 8:1, 97–108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>Michaela Mahlberg</td>
<td>The textlinguistic dimension of corpus linguistics: the support function of English general nouns and its theoretical implications</td>
<td>TESOL Quarterly 37:3, 513–527</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>Anna Mauranen</td>
<td>The corpus of English as lingua franca in academic settings</td>
<td>Research on Language and Social Interaction 36:1, 33–63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>Alan Partington</td>
<td>Footing shift for attribution: ‘according to the New York Times this morning . . .’</td>
<td>Logos and Language 4:2, 1–17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>Wolfgang Teubert</td>
<td>Writing, hermeneutics, and corpus linguistics</td>
<td>How to Use Corpora in Language Teaching, Amsterdam: Benjamins, pp. 15–36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>Mona Baker</td>
<td>A corpus-based view of similarity and difference in translation</td>
<td>System 30:2, 167–184</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>Silvia Bernardini</td>
<td>Corpora in the classroom: an overview and some reflections on future developments</td>
<td>International Journal of Corpus Linguistics 8:1, 97–108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Author</td>
<td>Title</td>
<td>Source</td>
<td>Vol.</td>
<td>Chap.</td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
<td>-------</td>
<td>--------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>Year</td>
<td>Author(s)</td>
<td>Title</td>
<td>Pages</td>
<td>Source</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------------------</td>
<td>----------------------------------------------------------------------</td>
<td>-------</td>
<td>--------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>Mark Davies</td>
<td>Advanced research on syntactic and semantic change with the Corpus del Español</td>
<td>V 99</td>
<td>C. Pusch, J. Kabatek and W. Raible (eds.), Romance Corpus Linguistics II: Corpora and Diachronic Linguistics, Tübingen: Gunter Narr Verlag, pp. 203–214</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>Ute Römer</td>
<td>Looking at looking: functions and contexts of progressives in spoken English and ‘school’ English</td>
<td>IV 60</td>
<td>A. Renouf and A. Kehoe (eds.), The Changing Face of Corpus Linguistics, Papers from the 24th International Conference on English Language Research on Computerized Corpora (ICAME), Amsterdam: Rodopi, pp. 231–242</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>Robert de Beaugrande</td>
<td>'Corporate bridges' twixt text and language</td>
<td>I 4</td>
<td><a href="http://www.beaugrande.com">www.beaugrande.com</a> (specially edited version for this publication), pp. 1–18</td>
<td></td>
</tr>
</tbody>
</table>
FOREWORD

This collection has been over two years in gestation. We started with an extremely ambitious vision, to cover not only the very early origins of corpus linguistics (in works such as that by Zipf, Herdan, and Firth) but also the recent global spread of corpus activities in many languages, undertaken for various different purposes and applications.

However, during the period of preparation, we were gradually forced to accept that, despite the seemingly generous allocation of space in these volumes, we would not be able to achieve anything like the breadth of coverage we had envisaged. We eventually realized that the most we could do would be to provide a sample of some important papers in the field, in the hope of stimulating the readers’ interest sufficiently to convince them to pursue the numerous related works mentioned in the references and bibliographies.