

Contents

Volume 1

I. Origin and history of corpus linguistics – corpus linguistics vis-à-vis other disciplines

1.	Charles F. Meyer, Pre-electronic corpora	1
2.	Fred Karlsson, Early generative linguistics and empirical methodology	14
3.	Stig Johansson, Some aspects of the development of corpus linguistics in the 1970s and 1980s	33
4.	Matti Rissanen, Corpus linguistics and historical linguistics	53
5.	Stefanie Dipper, Theory-driven and corpus-driven computational linguistics, and the use of corpora	68
6.	Suzanne Romaine, Corpus linguistics and sociolinguistics	96
7.	Ute Römer, Corpora and language teaching	112
8.	Ulrich Heid, Corpus linguistics and lexicography	131

II. Corpus compilation and corpus types

9.	Susan Hunston, Collection strategies and design decisions	154
10.	Marianne Hundt, Text corpora	168
11.	Anne Wichmann, Speech corpora and spoken corpora	187
12.	Jens Allwood, Multimodal corpora	207
13.	Joakim Nivre, Treebanks	225
14.	Claudia Claridge, Historical corpora	242
15.	Sylviane Granger, Learner corpora	259
16.	Karin Aijmer, Parallel and comparable corpora	275
17.	Michael Beißwenger/Angelika Storrer, Corpora of computer-mediated communication	292
18.	Gunnar Bergh/Eros Zanchetta, Web linguistics	309
19.	Alexander Mehler, Large text networks as an object of corpus linguistic studies	328

III. Existing corpora

20.	Richard Xiao, Well-known and influential corpora	383
21.	Nicholas Ostler, Corpora of less studied languages	457

IV. Preprocessing corpora

22.	Timm Lehmberg/Kai Wörner, Annotation standards	484
23.	Eric Atwell, Development of tag sets for part-of-speech tagging	501
24.	Helmut Schmid, Tokenizing and part-of-speech tagging	527
25.	Arne Fitschen/Piklu Gupta, Lemmatising and morphological tagging	552

26.	Paul Rayson/Mark Stevenson, Sense and semantic tagging	564
27.	Ruslan Mitkov, Corpora for anaphora and coreference resolution . . .	579
28.	Hannah Kermes, Syntactic preprocessing	598
29.	Dawn Archer/Jonathan Culpeper/Matthew Davies, Pragmatic annotation	613
30.	Nelleke Oostdijk/Lou Boves, Preprocessing speech corpora: Transcription and phonological annotation	642
31.	Peter Wittenburg, Preprocessing multimodal corpora	664
32.	Michael P. Oakes, Preprocessing multilingual corpora	685
33.	Martin Wynne, Searching and concordancing	706
34.	Sean Wallis, Searching treebanks and other structured corpora	738
35.	Heike Zinsmeister/Erhard Hinrichs/Sandra Kübler/Andreas Witt, Linguistically annotated corpora: Quality assurance, reusability and sustainability	759

Volume 2

V. Use and exploitation of corpora

36.	Marco Baroni/Stefan Evert, Statistical methods for corpus exploitation	777
37.	Marco Baroni, Distributions in text	803
38.	Douglas Biber, Multi-dimensional approaches	822
39.	Antal van den Bosch, Machine learning	855
40.	Hermann Moisl, Exploratory multivariate analysis	874
41.	R. Harald Baayen, Corpus linguistics in morphology: Morphological productivity	899
42.	W. Detmar Meurers/Stefan Müller, Corpora and syntax	920
43.	Anatol Stefanowitsch/Stefan Th. Gries, Corpora and grammar	933
44.	Sabine Schulte im Walde, The induction of verb frames and verb classes from corpora	952
45.	Michael Hoey, Corpus linguistics and word meaning	972
46.	Richard Xiao, Theory-driven corpus research: Using corpora to inform aspect theory	987
47.	Michael McCarthy/Anne O'Keeffe, Corpora and spoken language . . .	1008
48.	Anders Lindström/Robert Eklund, Cross-lingual influence: The integration of foreign items	1024
49.	Tuija Virtanen, Corpora and discourse analysis	1043
50.	Michael P. Oakes, Corpus linguistics and stylometry	1070
51.	Anne Curzan, Historical corpus linguistics and evidence of language change	1091
52.	Christian Mair, Corpora and the study of recent change in language . .	1109
53.	Lieselotte Anderwald/Benedikt Szmrecsanyi, Corpus linguistics and dialectology	1126
54.	Josef Schmied, Contrastive corpus studies	1140
55.	Silvia Hansen-Schirra/Elke Teich, Corpora in human translation	1159
56.	Harold Somers, Corpora and machine translation	1175

57.	Holger Diessel, Corpus linguistics and first language acquisition	1197
58.	Stefan Evert, Corpora and collocations	1212
59.	Paul Clough/Rob Gaizauskas, Corpora and text re-use	1249
60.	Constantin Orasan/Laura Hasler/Ruslan Mitkov, Corpora for text summarisation	1271
61.	Douglas Biber/James K. Jones, Quantitative methods in corpus linguistics	1287
Indexes (person index, corpus index, subject index)		1305

