

# Contents

Introduction . . . . .	1
<i>Stig Johansson</i> Computer corpora in English language research . . . . .	3
Probabilistic grammatical analysis . . . . .	7
<i>Steven J. DeRose</i> An analysis of probabilistic grammatical tagging methods . . . . .	9
<i>Geoffrey Leech and Roger Garside</i> Running a grammar factory: The production of syntactically analysed corpora or “treebanks” . . . . .	15
<i>Clive Souter and Tim F. O’Donoghue</i> Probabilistic parsing in the COMMUNAL project . . . . .	33
Syntax . . . . .	49
<i>Pieter de Haan</i> On the exploration of corpus data by means of problem-oriented tagging: Postmodifying clauses in the English noun phrase . . . . .	51
<i>Christian Mair</i> Quantitative or qualitative corpus analysis? Infinitival complement clauses in the Survey of English Usage corpus . . . . .	67
Lexis . . . . .	81
<i>Magnar Brekke</i> Automatic parsing meets the wall . . . . .	83
<i>Piek Vossen</i> Polysemy and vagueness of meaning descriptions in the <i>Longman dictionary of contemporary English</i> . . . . .	105

Speech . . . . .	125
<i>Bengt Altenberg</i>	
Amplifier collocations in spoken English . . . . .	127
<i>Gerry Knowles</i>	
Prosodic labelling: The problem of tone group boundaries . . . . .	149
<i>Anne Wichmann</i>	
A study of up-arrows in the Lancaster/IBM Spoken English Corpus	165
Regional/social variation . . . . .	179
<i>Peter Collins</i>	
<i>Will and shall</i> in Australian English . . . . .	181
<i>Ossi Ihalainen</i>	
The grammatical subject in educated and dialectal English: Comparing the London-Lund Corpus and the Helsinki Corpus of Modern English Dialects . . . . .	201
<i>Gerhard Leitner</i>	
The Kolhapur Corpus of Indian English – intravarietal description and/or intervarietal comparison . . . . .	215
Specialised corpora . . . . .	233
<i>Dorrit Faber and Karen M. Lauridsen</i>	
The compilation of a Danish-English-French corpus in contract law	235
<i>Magnus Ljung</i>	
Swedish TEFL meets reality . . . . .	245
Software . . . . .	257
<i>Benny Brodda</i>	
Doing corpus work with PC Beta; or, how to be your own computational linguist . . . . .	259
<i>Knut Hofland</i>	
Concordance programs for personal computers . . . . .	283
<i>Jacques Noël</i>	
Corpora and dictionaries in multiline records: A Unix/Awk approach to home-made text retrieval . . . . .	307

Reference section . . . . . 317

*Lita Taylor, Geoffrey Leech and Steven Fligelstone*

A survey of English machine-readable corpora . . . . . 319

*Bengt Altenberg*

A bibliography of publications relating to English computer corpora 355

Index . . . . . 397