Twentieth-Century English

Standard English has evolved and developed in many ways over the past hundred years. From pronunciation to vocabulary to grammar, this concise survey clearly documents the recent history of standard English. Drawing on large amounts of authentic corpus data, it shows how we can track ongoing changes to the language, and demonstrates each of the major developments that have taken place. As well as taking insights from a vast body of literature, Christian Mair presents the results of his own cutting-edge research, revealing some important changes which have not been previously documented. He concludes by exploring how social and cultural factors, such as the American influence on British English, have affected standard English in recent times. Authoritative, informative, and engaging, this book will be essential reading for anyone interested in language change in progress - particularly those working on English, and will be welcomed by students, researchers, and language teachers alike.

CHRISTIAN MAIR is Chair in English Linguistics at the Universität Freiburg, Germany, with research interests in the corpus-based description of modern English grammar, and in the study of regional variation and ongoing changes in standard English worldwide. He is author of Infinitival clauses in English: a study of syntax in discourse (Cambridge University Press, 1990).
STUDIES IN ENGLISH LANGUAGE

General editor: Merja Kyttö (Uppsala University)

Editorial Board: Bas Aarts (University College London), John Algeo (University of Georgia), Susan Fitzmaurice (Northern Arizona University), Richard Hogg (University of Manchester), Charles F. Meyer (University of Massachusetts)

Twentieth-Century English
STUDIES IN ENGLISH LANGUAGE

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 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 on 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 Paivi 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 evolution
Raymond Hickey (ed.) Legacies of colonial English
Merja Kyttö, Mats Rydén and Erik Smitterberg (eds.) Nineteenth century English: stability and change
John Algeo British or American English? A handbook of word and grammar patterns
Twentieth-Century English

History, Variation, and Standardization

CHRISTIAN MAIR
Contents

List of figures page ix
List of tables xi
Acknowledgments xv

1 Setting the scene 1

2 Ongoing language change: problems of detection and verification 12
2.1 “Visible” and “invisible” changes 12
2.2 The pitfalls of anecdotal observation 15
2.3 Documenting change 21
2.4 Outlook: a plea for methodological pluralism 33

3 Lexical change in twentieth-century English 36
3.1 Introduction 36
3.2 Case studies 38
3.3 Major trends 69
3.4 Neologizing in its social context

4 Grammatical changes in twentieth-century English 82
4.1 Introduction 82
4.2 Review of the literature 84
4.3 Aspect: twentieth-century changes in the structure and use of the progressive 88
4.4 The going to-future 95
4.5 Modality: must and shall – two modals on the way out, and possible replacements 100
4.6 Further developments in tense, aspect, modality: a synopsis of current research 108
4.7 Current changes in the English voice system 111
4.8 Nonfinite verb forms: some twentieth-century developments in the field of clausal complementation 119
### Contents

4.9 Nouns, pronouns, adjectives 140  
4.10 Conclusion 154

5 Pronunciation 156  
5.1 Introduction 156  
5.2 A history of RP in the twentieth century 162  
5.3 “General American”: myth or reality? 173

6 Language change in context: changing communicative and discourse norms in twentieth-century English 181  
6.1 Introduction 181  
6.2 The colloquialization of written English in the twentieth century 183  
6.3 Americanization? 193  
6.4 Analysis of selected sample texts 195

Conclusion 200

Appendix 1 Brief survey of the corpora used for the present study 206  
Appendix 2 The OED Baseline Corpora 210  
Appendix 3 Estimating text size in the newspaper archives and the World Wide Web 213  
Appendix 4 A quarterly update of the OED Online (New Edition) – 13 March 2003: Motswana to mussy 217

References 231  
Index 242
Figures

1.1 Languages of publication in five natural sciences (1879–1980), (Tsunoda 1983)  
2.1 Four matching one-million-word corpora of written English  
3.1 Frequency of use of selected computer neologisms in *The Guardian (and Observer) on CD-ROM*  
3.2 Frequency of use of selected military neologisms in *The Guardian (and Observer) on CD-ROM*  
3.3 March 2003 OED updates for words containing the combining form *multi-*  
3.4 March 2003 OED updates – out-of-sequence entries  
3.5 Frequency of selected verbs of the *up/down + V* type in *The Guardian (and Observer) on CD-ROM*  
3.6 Spread of three deverbal adjectives in *The Guardian (and Observer) on CD-ROM*  
4.1 *Going to* and *gonna* 1600–2000 – frequency as n/10,000 citations  
4.2 *Get*-passives according to age in the BNC  
4.3 Nonfinite complements of *remember* in the OED quotation base by century – normalized frequency as n/10,000 quotations  
4.4 *Help* + infinitive 1600–2000 – frequency as n/10,000 quotations  
4.5 Analytical and synthetic comparison for four classes of adjectives (Kytö and Romaine 1997: 344)  
5.1 The decline of /œə/ (John Wells, source: http://www.phon.ucl.ac.uk/home/wells/)  
5.2 Major dialect areas in the US based on the *Dictionary of American regional English* (DARE) (Carver 1987) and the *Phonological atlas of North America* (Labov et al. 2006)
x List of figures

(source: http://www.ling.upenn.edu/phono_atlas/NationalMap/NationalMap.html) 178

5.3 Merger of the vowels in *cot* and *caught* (source: William Labov, http://www.ling.upenn.edu/phono_atlas/ICSLP4.html#Heading2) 180

A3.1 Textual growth in *The Guardian (and Observer)* on CD-ROM 215

A3.2 *Deep breath* in top-level Web domains, December 2002 to March 2003 216
# Tables

1.1 Population of major urban centers in the English-using world

1.2 Percentage of languages in natural science publications, 1980 to 1996

1.3 Percentage of languages used in publications in the humanities, 1974 to 1995 (adapted from a graph in Ammon 1998a: 167)

2.1 Prepositions following *different* in four corpora

2.2 Prepositions following *different* in regionally stratified Web material (Google, 30 May 2004)

2.3 Proportion of *on/upon* in four corpora

2.4 Proportion of *on/upon* in three samples from the OED quotation base

2.5 Lexical items most characteristic of four groups of speakers in a corpus of spoken British English (compiled from Rayson et al. 1997)

2.6 Frequency of *wanna* in the BNC per age group (x/1,000,000 words)

3.1 *Wicked* – frequency in the spoken-demographic BNC per million words by age group

3.2 *Massive* – frequency in the spoken-demographic BNC per million words by age group

3.3 OED Online – new words first attested in the twentieth century

3.4 Nineteenth- and twentieth-century borrowings from Yiddish in the OED

3.5 Twentieth-century borrowings from Spanish first attested in the OED

3.6 Proper nouns consisting entirely of capital letters: comparison of frequency in LOB and F-LOB
List of tables

3.7 Prefixed verbs in up- and down- in the OED (twentieth century) and the BNC (compiled from information in Scheible 2005) 64
3.8 Up/down + V in four corpora – types (tokens) 64
3.9 Ongoing in four corpora 66
3.10 On/off/in/out/up/down + V + ing in four corpora – survey 67
3.11 On/off/in/out/up/down + V + ing in four corpora – type/token ratios 67
3.12 Discourse frequency of virtue(s), value(s), etc. in selected twentieth-century corpora 78
4.1 Progressive forms in the press sections (A–C) of four reference corpora 90
4.2 Been being and be being on the English-language Web (Google, 23 July 2003) 92
4.3 Going to-futures in four corpora (examples from direct speech in brackets) 96
4.4 Going to – manually post-edited output for four quarter-centuries 98
4.5 Will/shall- and going to-futures in four spoken corpora (percentages, adapted from Szmrecsanyi 2003: 303) 98
4.6 Going to- and will-futures in two age groups in the spoken-demographic BNC (normalized frequencies/words per million) 99
4.7 Use of the modal auxiliaries in four corpora 101
4.8 Shall-futures in four corpora 102
4.9 Have (got) to in four corpora (press texts, section A–C only) 104
4.10 Have got to in four corpora 104
4.11 Have to in four corpora 104
4.12 Obligation and necessity in the Santa Barbara Corpus of Spoken American English and the conversation component of ICE-GB 105
4.13 Must and have to by function in ICE-GB (spoken), adapted from Depraetere and Verhulst (forthcoming) 105
4.14 Get-passives in four corpora (examples from direct speech in brackets) 113
4.15 Frequency indices for get-passives in the BNC 116
4.16 For + NP + to-infinitival clauses in four corpora 124
4.17 For + NP + to-infinitival clauses in three OED Baseline Corpora 124
4.18 For + NP + to-infinitival clauses in three OED Baseline Corpora and F-LOB and Frown (normalized, as instances per million words) 125
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.19</td>
<td>Gerunds and infinitives after <em>remember</em> in the OED quotation base – normalized frequencies (“n/10,000 quotations,” rounded to the first decimal, absolute frequencies in brackets)</td>
</tr>
<tr>
<td>4.20</td>
<td>Notional subjects in gerundial constructions after <em>remember</em> in the OED quotation base</td>
</tr>
<tr>
<td>4.21</td>
<td>Proportion of infinitival and gerundial complements after <em>begin</em> in four corpora</td>
</tr>
<tr>
<td>4.22</td>
<td>Infinitive vs. gerund complements with <em>begin</em> in selected British databases</td>
</tr>
<tr>
<td>4.23</td>
<td>Infinitive vs. gerund complements with <em>begin</em> in selected American databases</td>
</tr>
<tr>
<td>4.24</td>
<td><em>To</em>-infinitive: V-ing after <em>start</em> in four corpora</td>
</tr>
<tr>
<td>4.25</td>
<td>Nonfinite complements of <em>prevent</em> in the OED Baseline Corpus</td>
</tr>
<tr>
<td>4.26</td>
<td>Ratio of <em>prevent</em> NP from V-ing vs. <em>prevent</em> NP V-ing in four corpora</td>
</tr>
<tr>
<td>4.27</td>
<td>Ratio of <em>stop</em> NP from V-ing vs. <em>stop</em> NP V-ing in four corpora</td>
</tr>
<tr>
<td>4.28</td>
<td><em>Save</em> NP V-ing in selected top-level Web domains</td>
</tr>
<tr>
<td>4.29</td>
<td><em>To-</em> vs. bare infinitives with <em>help</em> in four corpora</td>
</tr>
<tr>
<td>4.30</td>
<td>Complementation of <em>help</em> in the “spoken-demographic” BNC</td>
</tr>
<tr>
<td>4.31</td>
<td><em>Whom</em> in four matching corpora</td>
</tr>
<tr>
<td>4.32</td>
<td>Nominative vs. objective case for pronouns in specific syntactic contexts (BNC)</td>
</tr>
<tr>
<td>4.33</td>
<td><em>S</em>-genitives in four tagged corpora</td>
</tr>
<tr>
<td>4.34</td>
<td>Use of genitive and of-phrase in the press sections (A–C) of two corpora (compiled from Raab-Fischer 1995)</td>
</tr>
<tr>
<td>4.35</td>
<td>Normalized frequencies (occurrences per million words) for selected genitives in spoken and written text types from the BNC</td>
</tr>
<tr>
<td>4.36</td>
<td><em>Fewer</em> vs. <em>less</em> with count nouns in selected corpora</td>
</tr>
<tr>
<td>5.1</td>
<td>Variably pronounced words in contemporary RP (John Wells, Longman Pronunciation Survey)</td>
</tr>
<tr>
<td>6.1</td>
<td>Frequencies of <em>say</em> in selected genre categories of four corpora</td>
</tr>
<tr>
<td>6.2</td>
<td>Verb and negative contractions in the four corpora (from Leech and Smith 2005)</td>
</tr>
<tr>
<td>6.3</td>
<td>Contraction ratios (<em>not</em>-contractions) in journalistic and academic prose</td>
</tr>
<tr>
<td>6.4</td>
<td>Decline in frequency of use of the <em>be</em>-passive in the four reference corpora (from Leech and Smith 2005)</td>
</tr>
</tbody>
</table>
xiv List of tables

6.5 Frequency of noun+common-noun combinations in four corpora 192
A2.1 Composition of the OED Baseline Corpora 210
A2.2 The five most common words in Baseline, Brown, LOB, Frown and F-LOB 211
A2.3 “Absolute frequencies (rank)” for selected function words in five corpora 212
A3.1 Frequency of selected collocations in two corpora and eight newspaper archival discs 214
A3.2 Frequencies of ten collocations in the BNC and selected top-level Web domains 215
Acknowledgments

When I started working on the corpus-based investigation of change in progress in present-day English in the early 1990s, writing a book on this topic soon began to seem an attractive idea, though one which was bound to remain rather theoretical for a long time. Now that the book is about to be published, I would like to take the opportunity to thank a number of organizations and people without whose help it is unlikely that *Twentieth-Century English* would have seen the light of day a mere fifteen years after the idea for the book was first conceived.

To the *Deutsche Forschungsgemeinschaft* (DFG) I owe thanks for generously supporting two corpus-related research projects from 1994 to 1996 and from 2003 to 2006. Without their funding, F-LOB and Frown, two corpora providing important evidence for the present study, would have been completed much later (if at all), and various laborious but extremely useful annotation schemes and other enhancements would not even have been attempted. If the DFG gave the money, more members of my team at Freiburg than can be named here have given their expertise and dedication over the years. If I single out Marianne Hundt, Andrea Sand, Stefanie Rapp, Birgit Waibel, and Lars Hinrichs by name, I hope that many others involved in the projects for longer or shorter periods of time will not take this amiss.

At CUP, I would like to thank Kate Brett, who, after discussing the idea of a history of twentieth-century English with me at the Edinburgh Late Modern English conference in 1998, encouraged me to formalize it by submitting a proposal to the Press. Helen Barton, who eventually took over from Kate, was equally sympathetic and additionally showed welcome patience in the final stages of completing the manuscript. Valuable suggestions for improvements were made by Merja Kytö, one of the series editors, who carefully went through the first version of the completed manuscript. At the very end of the production process, working together with Nikky Twyman as a copy editor was a pleasant and humbling experience, pleasant because of her quiet and good-humored efficiency, and humbling because of the number of oversights she spotted in a manuscript which I thought I had proofread carefully.
xvi  Acknowledgments

I hope that the book will convey to its readers some of my own fascination with the “living history” of English, its recent past, its rich and diversified present, and its future, and that it will encourage others to keep researching the many questions which I have had to leave unanswered.

Freiburg, February 2006         CM