

Cambridge University Press

978-1-107-03247-7 - Regional Variation in Written American English

Jack Grieve

Frontmatter

[More information](#)

REGIONAL VARIATION IN WRITTEN AMERICAN ENGLISH

The first study of its kind, *Regional Variation in Written American English* takes a corpus-based approach to map over 100 grammatical alternation variables across the United States. A multivariate spatial analysis of these maps shows that grammatical alternation variables follow a relatively small number of common regional patterns in American English, which can be explained based on both linguistic and extra-linguistic factors. Through this rigorous analysis of extensive data, Grieve identifies five primary modern American dialect regions, demonstrating that regional variation is far more pervasive and complex in natural language than is generally assumed. The wealth of maps and data and the ground-breaking implications of this volume make it essential reading for students and researchers in linguistics, English language, geography, computer science, sociology, and communication studies.

- Identifies and maps regional linguistic variation in written Standard English for the first time.
- Introduces a corpus-based approach to dialectology.
- Presents a statistical method for identifying individual and common patterns of regional variation.

JACK GRIEVE is Senior Lecturer in Forensic Linguistics in the School of Languages and Social Sciences at Aston University in Birmingham, England. He holds a Ph.D. in Applied Linguistics from Northern Arizona University, where he studied quantitative corpus linguistics under the supervision of Professor Douglas Biber. He was also a postdoctoral research fellow in Professor Dirk Geeraerts's Quantitative Lexicology and Variational Linguistics research unit at the University of Leuven in Belgium.

Cambridge University Press
 978-1-107-03247-7 - Regional Variation in Written American English
 Jack Grieve
 Frontmatter
[More information](#)

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:

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*

Martin Hilpert: *Constructional Change in English: Developments in allomorphy, word formation, and syntax*

Jakob R. E. Leimgruber: *Singapore English: Structure, Variation and Usage*

Cambridge University Press
978-1-107-03247-7 - Regional Variation in Written American English
Jack Grieve
Frontmatter
[More information](#)

- Christoph Rühlemann: *Narrative in English Conversation*
Dagmar Deuber: *English in the Caribbean: Variation, Style and Standards in Jamaica and Trinidad*
Eva Berlage: *Noun Phrase Complexity in English*
Nicole Dehé: *Parentheticals in Spoken English: The Syntax-Prosody Relation*
Jock Onn Wong: *English in Singapore: A Cultural Analysis*
Anita Auer, Daniel Schreier and Richard J. Watts: *Letter Writing and Language Change*
Marianne Hundt: *Late Modern English Syntax*
Irma Taavitsainen, Merja Kyto, Claudia Claridge, and Jeremy Smith: *Developments in English: Expanding Electronic Evidence*
Arne Lohmann: *English Co-ordinate Constructions: A Processing Perspective on Constituent Order*
John Flowerdew and Richard W. Forest: *Signalling Nouns in English: A corpus-based discourse approach*
Jeffrey P. Williams, Edgar W. Schneider, Peter Trudgill, and Daniel Schreier: *Further Studies in the Lesser-Known Varieties of English*
Nuria Yáñez-Bouza: *Grammar, Rhetoric and Usage in English: Preposition Placement 1500–1900*
Jack Grieve: *Regional Variation in Written American English*

Earlier titles not listed are also available

Cambridge University Press

978-1-107-03247-7 - Regional Variation in Written American English

Jack Grieve

Frontmatter

[More information](#)

Cambridge University Press

978-1-107-03247-7 - Regional Variation in Written American English

Jack Grieve

Frontmatter

[More information](#)

REGIONAL VARIATION IN WRITTEN AMERICAN ENGLISH

JACK GRIEVE

Aston University



CAMBRIDGE
UNIVERSITY PRESS

Cambridge University Press
978-1-107-03247-7 - Regional Variation in Written American English
Jack Grieve
Frontmatter
[More information](#)

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.
It furthers the University's mission by disseminating knowledge in the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107032477

© Jack Grieve 2016

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2016

A catalogue record for this publication is available from the British Library

ISBN 978-1-107-03247-7 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy
of URLs for external or third-party internet websites referred to in this publication,
and does not guarantee that any content on such websites is, or will remain,
accurate or appropriate.

Cambridge University Press

978-1-107-03247-7 - Regional Variation in Written American English

Jack Grieve

Frontmatter

[More information](#)

For Emily

Cambridge University Press

978-1-107-03247-7 - Regional Variation in Written American English

Jack Grieve

Frontmatter

[More information](#)

Contents

<i>List of figures</i>	<i>page</i> xii
<i>List of tables</i>	xv
<i>Acknowledgements</i>	xvii
1 Introduction	1
1.1 American dialectology	1
1.2 Outline	13
2 Corpus	16
2.1 Corpus-based dialectology	16
2.2 Corpus design	19
2.2.1 Register selection	19
2.2.2 City selection	23
2.3 Corpus compilation	27
2.4 Corpus dimensions	29
3 Grammatical analysis	36
3.1 Alternation variables	36
3.2 Alternation variable selection and measurement	39
3.3 Alternation variable mapping	43
3.4 Feature set	47
3.4.1 Pronouns	50
3.4.2 Relative pronouns	52
3.4.3 Determiners	52
3.4.4 Adjectives	54
3.4.5 Nouns	57
3.4.6 Prepositions	58
3.4.7 Particles	70
3.4.8 Subordinators	72
3.4.9 Coordinators	74
3.4.10 Verbs	76
3.4.11 Modals	78
3.4.12 Infinitives	81

x	<i>Contents</i>	
	3.4.13 Not	82
	3.4.14 Adverbs	83
	3.4.15 Adverbials	87
4	Spatial analysis	98
	4.1 Standard approaches	98
	4.2 Regional linguistic data matrix	102
	4.3 Introduction to spatial analysis	103
	4.4 Local spatial autocorrelation analysis	107
	4.5 Spatial weights matrix	114
	4.6 Local spatial autocorrelation maps	122
	4.7 Discussion of results	122
5	Multivariate analysis	144
	5.1 Dialectometry	144
	5.2 Multivariate spatial analysis	147
	5.3 Spatially autocorrelated linguistic data matrix	150
	5.3.1 Alternation variable selection	150
	5.3.2 Missing data	154
	5.4 Factor analysis	156
	5.4.1 Introduction to factor analysis	156
	5.4.2 Preliminary analyses	161
	5.4.3 Results	166
	5.5 Fuzzy cluster analysis	175
	5.5.1 A note on dialect regions	175
	5.5.2 Introduction to fuzzy cluster analysis	176
	5.5.3 Results	178
6	Sources of regional linguistic variation	185
	6.1 Internal explanations	185
	6.1.1 Factor 1	185
	6.1.2 Factor 2	189
	6.1.3 Factor 3	193
	6.1.4 Regional variation in formality	195
	6.2 External explanations	198
	6.3 Comparison of internal and external explanations	207
	6.4 Comparison to previous American dialect surveys	209
	6.5 Future American dialect regions	216
7	Conclusion	219
	<i>Appendices</i>	
	<i>Appendix A: Corpus dimensions</i>	227
	<i>Appendix B: Grammatical alternation variables: search patterns</i>	255

Cambridge University Press

978-1-107-03247-7 - Regional Variation in Written American English

Jack Grieve

Frontmatter

[More information](#)

<i>Contents</i>	xi
<i>Appendix C: Grammatical alternation variables: descriptive statistics</i>	288
<i>Appendix D: Spatial analysis results</i>	299
<i>Appendix E: Multivariate results</i>	311
<i>References</i>	320
<i>Index</i>	330

Figures

1.1	American dialect regions: <i>Linguistic Atlas of United States and Canada</i>	page 4
1.2	American dialect regions: <i>Dictionary of American English</i>	10
1.3	American dialect regions: <i>Atlas of North American English</i>	12
2.1	City sub-corpora	32
2.2	Most common year of publication	33
2.3	Words per sub-corpus	34
2.4	Letters per sub-corpus	34
2.5	Authors per sub-corpus	35
3.1	<i>Be Not</i> contraction	43
3.2	<i>Be Not</i> full	44
3.3	<i>Although</i> alternation	45
3.4	<i>Though</i> alternation	45
3.5	<i>Even though</i> alternation	46
3.6	<i>Be Not</i> contraction simplified maps	46
3.7	<i>Although, Though, and Even Though</i> simplified maps	47
3.8	Pronoun alternation maps	51
3.9	Relative pronoun alternation maps	54
3.10	Determiner alternation maps (Set 1)	56
3.11	Determiner alternation maps (Set 2)	58
3.12	Adjective alternation maps	60
3.13	Noun alternation maps	62
3.14	Preposition alteration maps (Set 1)	63
3.15	Preposition alteration maps (Set 2)	65
3.16	Preposition alteration maps (Set 3)	67
3.17	Preposition alteration maps (Set 4)	69
3.18	Particle alternation maps	71
3.19	Subordinator alternation maps	73
3.20	Coordinator alternation maps	75
3.21	Verb alternation maps	77

List of figures

xiii

3.22	Modal alternation maps	80
3.23	Infinitive alternation maps	82
3.24	Not alternation maps	84
3.25	Adverb alternation maps (Set 1)	86
3.26	Adverb alternation maps (Set 2)	88
3.27	Adverbial alternation maps (Set 1)	90
3.28	Adverbial alternation maps (Set 2)	93
3.29	Adverbial alternation maps (Set 3)	95
3.30	Adverbial position alternation maps	97
4.1	Histograms for the variants of selected alternations	104
4.2	<i>Anyone/Anybody</i> map	111
4.3	<i>Anyone/Anybody</i> local spatial autocorrelation map	112
4.4	<i>Anyone</i> and <i>Anybody</i> simplified local spatial autocorrelation map	112
4.5	Comparison of spatial weights matrices for <i>Anyone/Anybody</i>	119
4.6	Comparison of spatial weights matrices for <i>Do Not</i> contraction	120
4.7	Local spatial autocorrelation maps	123
4.8	MAG_i score by word class box plots	143
5.1	Interpolated local spatial autocorrelation map for <i>Rarely/Seldom</i>	155
5.2	Additional amount of variance explained chart and scree plot	163
5.3	Factor 1 map	170
5.4	Factor 2 map	170
5.5	Factor 3 map	171
5.6	Factor 1 vs. Factor 2 cities	171
5.7	Factor 1 vs. Factor 2 states	172
5.8	Factor 1 vs. Factor 3 cities	172
5.9	Factor 1 vs. Factor 3 states	173
5.10	Factor 2 vs. Factor 3 cities	173
5.11	Factor 2 vs. Factor 3 states	174
5.12	Northeastern dialect region cluster	179
5.13	Midwestern dialect region cluster	180
5.14	Southeastern dialect region cluster	180
5.15	South Central dialect region cluster	181
5.16	Western dialect region fuzzy cluster	181
5.17	Overall dialect regions	182
5.18	Factor 1 vs. Factor 2 dialect regions	183
5.19	Factor 1 vs. Factor 3 dialect regions	183
5.20	Factor 2 vs. Factor 3 dialect regions	184

Cambridge University Press

978-1-107-03247-7 - Regional Variation in Written American English

Jack Grieve

Frontmatter

[More information](#)

xiv

List of figures

6.1	Average formality factor scores (based on Factors 1, 2, 3)	198
6.2	Elevation	200
6.3	Population density (2011)	201
6.4	Per capita income by county (1999)	202
6.5	Percentage Republican presidential vote (2008)	203
6.6	Percentage German ancestry	204
6.7	Percentage African American (2010)	204
6.8	Census Bureau regions	205
6.9	Modern American dialect regions	210
6.10	Phonetic Factor 1 (<i>Atlas of North American English</i> reanalysis)	212
6.11	Phonetic Factor 2 (<i>Atlas of North American English</i> reanalysis)	213
6.12	Phonetic Factor 3 (<i>Atlas of North American English</i> reanalysis)	213
6.13	Phonetic Factor 4 (<i>Atlas of North American English</i> reanalysis)	214
6.14	Phonetic dialect regions (<i>Atlas of North American English</i> reanalysis)	214
6.15	Future American dialect regions	218

Tables

2.1	Letter to the editor examples	<i>page</i> 21
2.2	City sub-corpora	30
3.1	The complete set of 135 grammatical alternations and their 295 variants	48
3.2	Pronoun alterations	50
3.3	English relative pronoun system	52
3.4	Relative pronoun alterations	53
3.5	Determiner alterations (Set 1)	55
3.6	Determiner alterations (Set 2)	57
3.7	Adjective alterations	59
3.8	Noun alterations	61
3.9	Preposition alterations (Set 1)	63
3.10	Preposition alterations (Set 2)	64
3.11	Preposition alterations (Set 3)	66
3.12	Preposition alterations (Set 4)	68
3.13	Particle alternations	70
3.14	Subordinator alternations	72
3.15	Coordinator alternations	74
3.16	Verb alternations	76
3.17	Modal alternations	79
3.18	Infinitive alternations	81
3.19	<i>Not</i> alternations	83
3.20	Adverbs alternations (Set 1)	85
3.21	Adverbs alternations (Set 2)	87
3.22	Adverbial alternations (Set 1)	89
3.23	Adverbial alternations (Set 2)	92
3.24	Adverbial alternations (Set 3)	94
3.25	Adverbial position alternations	96
4.1	Structure of the regional linguistic data matrix	103
4.2	Top 25 alternations/variants by <i>MAG_i</i> score	142

Cambridge University Press

978-1-107-03247-7 - Regional Variation in Written American English

Jack Grieve

Frontmatter

[More information](#)

xvi

List of tables

5.1	Preliminary factor analyses	162
5.2	Factor loadings and variable communalities	167
6.1	Factor 1 variant loadings	186
6.2	Factor 2 variant loadings	190
6.3	Relative frequency of selected alternations PMW in COCA (Factor 2)	191
6.4	Factor 3 variant loadings	193

Acknowledgements

I began working on the research reported in this book ten years ago as part of my Ph.D. dissertation. I had finished my M.A. at Simon Fraser University in Vancouver with Paul McFetridge and Maite Taboada, and moved to Northern Arizona University in Flagstaff to work with Doug Biber. I was incredibly lucky to be supervised by Doug, who more than anyone else is responsible for where I am today as an academic. His influence can be seen throughout this book, both in the methods I have applied and the manner in which I have interpreted my results. I am also grateful for the support and guidance I received from the members of my doctoral committee, Bill Crawford, Ray Huang, and Randi Reppen.

Following my Ph.D., I became a researcher at the Quantitative Lexicology and Variational Linguistics (QLVL) research unit at the University of Leuven, headed by Dirk Geeraerts and Dirk Speelman, whose theoretical and methodological outlook on linguistics has also shaped this book. While at QLVL, I worked closely with Costanza Asnaghi and Tom Ruetten. Many of the ideas and methods described in this book were developed during conversations with them, as well as with other members of QLVL, including Alena Anishchanka, Dirk De Hertog, Kris Heylen, Natalia Levshina, Pedro Pulquério Vieira, Eline Zenner, and Weiwei Zhang.

After Leuven, I took up a lectureship at Aston University in Birmingham as a member of the Centre for Forensic Linguistics (CFL), which is where this book was written. I am particularly grateful to Tim Grant, the head of CFL, for his constant support, as well as my other colleagues at CFL, including Kate Haworth, Krzys Kredens, and Andrea Nini, as well as Malcolm Coulthard, who founded the unit. I would also like to thank all my colleagues in the School of Languages and Social Sciences, especially Judith Baxter, Ursula Clark, Simon Green, Pam Moores, Garry Plappert, and Gertrud Reershemius.

The research reported in this book has also benefited from discussions with a number of linguists and social scientists from around the world,

Cambridge University Press

978-1-107-03247-7 - Regional Variation in Written American English

Jack Grieve

Frontmatter

[More information](#)

xviii

Acknowledgements

especially the dialectometrists Wilbert Heeringa, John Nerbonne, Benedikt Szmrecsanyi, and Martijn Weeling, as well as with Federica Barbieri, Eric Friginal, Diansheng Guo, Roland Van Hout, Dan Johnson, Alice Kasakoff, Yuri Thierein, and Bert Vaux. I would also like to thank Mark Davies and Bill Labov for sharing their data with me, as well as their thoughts on this project.

This book would also not have been possible without the support of Cambridge University Press, especially Helen Barton and Merja Kytö. I am grateful for the confidence they have had in this project and their feedback throughout this process has been invaluable. I would also like to thank Helena Dowson, Bethany Gaunt, Christina Sarigiannidou, and Martin Barr for their help preparing this book for publication.

Finally, I would like to thank my family, especially my sister, Meg, my father, Tom, and my mother, Paula, who passed away while I was studying for my Ph.D. I would also like to thank my father's partner Linda and my parents-in-law Bob and Kathleen for their love and support. Most of all, I must thank my wife, Emily, to whom I dedicate this book. I met Emily in Flagstaff and we were married in Leuven. She has read every line of this book, most several times, and we have discussed every part of this research together. This book would not have been possible without her.