

## Register Variation Online

While other books focus on special online registers, such as tweets or texting, no previous study describes the full range of everyday registers found on the searchable web. These registers are the documents that readers encounter every time they do a Google search; registers like news reports, product reviews, travel blogs, discussion forums, FAQs, etc. Based on analysis of a large, near-random corpus of web documents, this monograph provides comprehensive situational, lexical, and grammatical descriptions of those registers. Beginning with a coding of each document in the corpus, the description identifies the registers that are especially common on the searchable web versus those that are less commonly found. Multidimensional analysis is used to describe the overall patterns of linguistic variation among web registers, while the second half of the book provides an in-depth description of each individual register, including analyses of situational contexts and communicative purposes, together with the typical lexical and grammatical characteristics associated with those contexts.

**Douglas Biber** is Regents' Professor of English (Applied Linguistics) at Northern Arizona University. His research efforts have focused on corpus linguistics, English grammar, and register variation (in English and cross-linguistic; synchronic and diachronic). He has published over 220 research articles and twenty-three books and monographs, including primary research studies as well as textbooks.

**Jesse Egbert** is Assistant Professor in the Applied Linguistics program at Northern Arizona University. He specializes in corpus-based research on register variation, particularly academic writing and online language, and methodological issues in quantitative linguistic research. He has published more than thirty research articles in journals such as the *Journal of English Linguistics*, the *International Journal of Corpus Linguistics*, and *Corpus Linguistics and Linguistic Theory*.

Cambridge University Press  
978-1-107-12216-1 — Register Variation Online  
Douglas Biber , Jesse Egbert  
Frontmatter  
[More Information](#)

---

## Register Variation Online

---

Douglas Biber

*Northern Arizona University*

Jesse Egbert

*Northern Arizona University*



CAMBRIDGE  
UNIVERSITY PRESS

Cambridge University Press  
978-1-107-12216-1 — Register Variation Online  
Douglas Biber, Jesse Egbert  
Frontmatter  
[More Information](#)

## CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom  
One Liberty Plaza, 20th Floor, New York, NY 10006, USA  
477 Williamstown Road, Port Melbourne, VIC 3207, Australia  
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,  
New Delhi – 110025, India  
79 Anson Road, #06-04/06, Singapore 079906

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](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9781107122161](http://www.cambridge.org/9781107122161)

DOI: 10.1017/9781316388228

© Douglas Biber and Jesse Egbert 2018

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 2018

Printed and bound in Great Britain by Clays Ltd, Elcograf S.p.A.

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

*Library of Congress Cataloging-in-Publication Data*

Names: Biber, Douglas, author. | Egbert, Jesse, 1985- author.

Title: Register variation online / Douglas Biber, Jesse Egbert.

Description: Cambridge, United Kingdom ; New York, NY : Cambridge University Press, 2018. | Includes bibliographical references and index.

Identifiers: LCCN 2017060046 | ISBN 9781107122161 (hardback)

Subjects: LCSH: Register (Linguistics) | Context (Linguistics) | Language and the Internet.

Classification: LCC P302.815 .B537 2018 | DDC 306.44—dc23

LC record available at <https://lcn.loc.gov/2017060046>

ISBN 978-1-107-12216-1 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.

## Contents

---

<i>List of Figures and Screenshots</i>	<i>page</i> vii
<i>List of Tables</i>	ix
1 Introduction	1
2 Corpus and Methods	12
3 A Survey of the Registers on the Publicly Searchable Web	26
4 Overall Patterns of Register Variation on the Searchable Web: A Multidimensional Analysis	44
5 Narrative Registers	73
6 Opinion, Advice, and Persuasion Registers	106
7 Informational Descriptions, Explanations, and Procedures	135
8 Oral Registers	177
9 The Web as a Continuous Space of Register Variation	196
<i>Appendix A Linguistic Features Included in the Multidimensional Analysis</i>	217
<i>Appendix B Lexico-grammatical Features Included in the Key Feature Analysis</i>	219
<i>Appendix C Descriptive Statistics for the Key Feature Analyses</i>	221
<i>References</i>	243
<i>Index</i>	247

Cambridge University Press  
978-1-107-12216-1 — Register Variation Online  
Douglas Biber , Jesse Egbert  
Frontmatter  
[More Information](#)

---

## Figures and Screenshots

---

### Figures

2.1	Screenshot of the first page of the Google Survey instrument	<i>page</i> 18
2.2	Online interface for CORE	19
2.3	The frequency of the top fifteen most frequent pronouns in CORE	20
2.4	The frequency of pronouns in the first twenty registers in CORE	20
2.5	The first twenty-five concordance lines for pronouns within the Personal Blogs register	21
5.1	Dimension scores across the seven Narrative sub-register categories	81
5.2	Key grammatical features for the News Reports sub-register	83
5.3	Key grammatical features for the Personal Blogs sub-register	87
5.4	Key grammatical features for the Sports Reports sub-register	90
5.5	Key grammatical features for the Historical Articles sub-register	95
5.6	Key grammatical features for the Travel Blogs sub-register	99
5.7	Key grammatical features for the Short Stories sub-register	102
6.1	MD analysis results for the Opinion sub-registers	116
6.2	Key grammatical features for the Opinion Blogs	120
6.3	Key grammatical features for Reviews	123
6.4	Key grammatical features for the Description-with-intent-to-sell sub-register	125
6.5	Key grammatical features for the Advice sub-register	128
6.6	Key grammatical features for Religious Blogs	132
7.1	MD analysis results for the Informational sub-registers	150
7.2	Key grammatical features for the How-to/Instructional sub-register	155
7.3	Key grammatical features for Recipes	157
7.4	Key grammatical features for Research Articles	159
7.5	Key grammatical features for Encyclopedia Articles	162
7.6	Key grammatical features for Descriptions-of-a-person	165
7.7	Key grammatical features for FAQs	167

viii	List of Figures and Screenshots	
7.8	Key grammatical features for Informational Blogs	170
7.9	Key grammatical features for Other Informational documents	173
8.1	Dimension scores across the four Oral sub-register categories	183
8.2	Key grammatical features for the Interactive Discussion sub-register	186
8.3	Key grammatical features for the Lyrical sub-register	189
8.4	Key grammatical features for the Interview sub-register	193
9.1	Hierarchical cluster tree for the sub-register categories	199
9.2	Dimension 3 scores for selected registers and hybrids: oral narrative versus written information	206
9.3	Dimension 1 scores for selected registers and hybrids: oral-involved versus literate-informational	207
9.4	Dimension 5 scores for selected registers and hybrids: irrealis versus informational narration	207
9.5	Hierarchical cluster tree	210
<b>Screenshots</b>		
5.1	A Personal Blog post	76
5.2	Travel Blog	77
5.3	Short Story	78
6.1	Opinion Blog from a personal site	109
6.2	Product review from an institutional site	111
6.3	An online Advice page	114
7.1	Online How-to document	139
7.2	Online How-to document	140
7.3	Wikipedia article	142
7.4	Encyclopedia Article	143
7.5	A biographical description of Charles Dickens	144
7.6	An online biography	145
7.7	Informational Blog from a commercial website	147
8.1	Interactive Discussion	180
8.2	Online Poem	181
8.3	Interview	182



## Tables

---

1.1	Growth of the World Wide Web; 2000–2017	<i>page</i> 2
1.2	Growth of the World Wide Web expressed as proportional increases; 2000–2017	2
1.3	Register categories for the first five pages of hits (fifty-eight documents) returned by a Google search on the word <i>horse</i>	5
2.1	Visual representation of the key situational distinctions made in the final register framework	17
3.1	Agreement results for the general register classification of 48,571 web documents	28
3.2	Agreement results for the specific sub-register classification of 48,571 web documents	29
3.3	Frequency information for general register categories	30
3.4	Frequency information for narrative sub-register categories	32
3.5	Frequency information for informational sub-register categories	34
3.6	Frequency information for opinion sub-register categories	35
3.7	Frequency information for other sub-register categories	37
3.8	Fleiss' Kappa coefficients indicating the extent to which raters agreed in their perceptions of each register category	38
3.9	Frequency information for texts containing reader comments, by register (excluding interactive discussions)	39
3.10	Register categorization used for the linguistic descriptions in Chapters 4–8	42
4.1	Dimensions as predictors of general register variation (eight general register categories + “other”)	48
4.2	Dimensions as predictors of specific sub-register variation (twenty-five specific sub-registers)	48
4.3	Summary of results from the MD analysis	50
5.1	Sub-register categories in the Narrative category with text counts	73
5.2	Situational characteristics of Narrative sub-registers	75
5.3	Dimension score means and standard deviations for Narrative sub-registers	82
5.4	Keywords for the News Report sub-register	85
		ix

x	List of Tables	
5.5	Keywords for the Personal Blogs sub-register	88
5.6	Keywords for the Sports Report sub-register	92
5.7	Keywords for the Historical Articles sub-register	97
5.8	Keywords for the Travel Blogs sub-register	100
5.9	Keywords for the Short Stories sub-register	104
6.1	Sub-register categories in the Opinion/Advice/Persuasion category, with text counts	106
6.2	Major situational characteristics of Opinion, Advice, and Persuasion sub-registers	108
6.3	Example topical domains and postings for Opinion Blogs	110
6.4	Dimension score means and standard deviations for Opinion sub-registers	119
6.5	Keywords for Opinion Blogs	121
6.6	Keywords for Reviews	124
6.7	Keywords for the Description-with-intent-to-sell sub-register	127
6.8	Keywords for the Advice sub-register	130
6.9	Keywords for Religious Blogs	133
7.1	Sub-register categories related to informational procedures/ descriptions, with text counts	135
7.2	Major situational characteristics of informational sub-registers	137
7.3	Dimension score means and standard deviations for Informational sub-registers	153
7.4	How-to keywords	156
7.5	Recipes keywords	158
7.6	Research article keywords	161
7.7	Keywords for Encyclopedia Articles	163
7.8	Keywords for Descriptions-of-a-person	166
7.9	FAQs keywords	169
7.10	Informational Blogs keywords	172
7.11	Other Information keywords	174
8.1	Sub-register categories in the Oral category with text counts	178
8.2	Situational characteristics of Oral sub-registers	179
8.3	Dimension score means and standard deviations for Oral sub-registers	185
8.4	Keywords for the Interactive Discussion sub-register	188
8.5	Keywords for the Lyrical sub-register	191
8.6	Keywords for the Interview sub-register	194
9.1	General register 2+2 hybrid combinations (occurring more than one hundred times in the corpus)	201
9.2	General register 2+1+1 hybrid combinations (occurring more than one hundred times in the corpus)	202

List of Tables	xi
9.3 Specific sub-register 2+2 hybrid combinations (occurring more than one hundred times in the corpus)	204
9.4 Specific sub-register 2+1+1 hybrid combinations (occurring more than one hundred times in the corpus)	205
9.5 Summary of two-, three-, and four-cluster solutions	211
C.1 Key grammatical features for the News Report sub-register	221
C.2 Key grammatical features for the Personal Blogs sub-register	222
C.3 Key grammatical features for the Sports Report sub-register	223
C.4 Key grammatical features for the Historical Articles sub-register	224
C.5 Key grammatical features for the Travel Blogs sub-register	225
C.6 Key grammatical features for the Short Stories sub-register	226
C.7 Key grammatical features for the Opinion Blogs	227
C.8 Key grammatical features for Reviews	227
C.9 Key grammatical features for the Description-with-intent- to-sell sub-register	228
C.10 Key grammatical features for the Advice sub-register	229
C.11 Key grammatical features for Religious Blogs	230
C.12 Key grammatical features for the How-to/Instructional sub-register	231
C.13 Key grammatical features for Recipes	232
C.14 Key grammatical features for Research Articles	233
C.15 Key grammatical features for Encyclopedia Articles	234
C.16 Key grammatical features for Descriptions-of-a-person	235
C.17 Key grammatical features for FAQs	236
C.18 Key grammatical features for Informational Blogs	237
C.19 Key grammatical features for the Other Informational sub-register	238
C.20 Key grammatical features for the Interactive Discussion sub-register	239
C.21 Key grammatical features for the Lyrical sub-register	240
C.22 Key grammatical features for the Interview sub-register	241

Cambridge University Press  
978-1-107-12216-1 — Register Variation Online  
Douglas Biber , Jesse Egbert  
Frontmatter  
[More Information](#)

---