

## Discourse-Pragmatic Variation and Change

Discourse-pragmatic markers are central to everyday language, yet many aspects of their use and functions remain elusive or under-investigated. Bringing together a global team of leading scholars, this volume presents a representative showcase of work currently being conducted in the field of discourse-pragmatic variation and change, including investigations of features such as *uh/um*, *please*, sentence-final *is all*, and discourse-pragmatic features from a number of languages. The book emphasizes not only that researchers have answered the call to address complex issues such as cross-linguistic reliability, extending research across languages, and expanding and improving on methods and analysis, but that they continue to address perennial questions in the field of language variation and change. With sections on theoretical and methodological issues, innovative variables, and language contact situations, the volume offers a robust overview of best practices for both new and experienced researchers.

Elizabeth Peterson is a sociolinguist based in Finland. She is the author of numerous articles on pragmatic borrowing in Finnish and the 2020 book *Making Sense of "Bad English"*.

Turo Hiltunen works as a university lecturer in English at the University of Helsinki, Finland. He has authored several studies on corpus pragmatics, corpus compilation, phraseology, and the language of science and medicine.

Joseph Kern is an assistant professor in Spanish at the University of Virginia's College at Wise, USA. He is the author of several articles on discourse-pragmatic features in Spanish, especially in varieties of Spanish in contact.

# Discourse-Pragmatic Variation and Change

*Theory, Innovations, Contact*

---

*Edited by*

Elizabeth Peterson

*University of Helsinki*

Turo Hiltunen

*University of Helsinki*

Joseph Kern

*University of Virginia's College at Wise*



CAMBRIDGE  
UNIVERSITY PRESS

Cambridge University Press & Assessment

978-1-108-83620-3 — Discourse-Pragmatic Variation and Change: Theory, Innovations, Contact

Edited by Elizabeth Peterson, Turo Hiltunen, Joseph Kern

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

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

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/9781108836203](http://www.cambridge.org/9781108836203)

DOI: 10.1017/9781108864183

© Cambridge University Press 2022

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 2022

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

*Library of Congress Cataloging-in-Publication Data*

Names: Peterson, Elizabeth (Professor of linguistics), editor. | Hiltunen, Turo, editor. | Kern, Joseph, (Professor of Spanish), editor.

Title: Discourse-pragmatic variation and change : theory, innovations, contact / edited by Elizabeth Peterson, Turo Hiltunen, Joseph Kern.

Description: Cambridge, United Kingdom ; New York : Cambridge University Press, 2021. | Includes bibliographical references and index.

Identifiers: LCCN 2021026770 (print) | LCCN 2021026771 (ebook) |

ISBN 9781108836203 (hardback) | ISBN 9781108864183 (ebook)

Subjects: LCSH: Discourse markers. | Pragmatics. | Language and languages – Variation. | Languages in contact. | BISAC: LANGUAGE ARTS & DISCIPLINES / Linguistics / Semantics | LCGFT: Essays.

Classification: LCC P302.35 .D57 2021 (print) | LCC P302.35 (ebook) | DDC 401/.41–dc23/eng/20210930

LC record available at <https://lcn.loc.gov/2021026770>

LC ebook record available at <https://lcn.loc.gov/2021026771>

ISBN 978-1-108-83620-3 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</i>	page vii
<i>List of Tables</i>	ix
<i>Notes on Contributors</i>	xiii
<i>Foreword by Jan-Ola Östman</i>	xix
<i>Acknowledgments</i>	xxvi
<i>List of Abbreviations</i>	xxvii
Introduction	1
ELIZABETH PETERSON, TURO HILTUNEN, AND JOSEPH KERN	
<b>Part I Innovations in Theory and Method</b>	13
1 Reflexes of Abruptness in the Development of Pragmatic Markers	15
DEREK DENIS	
2 Evaluation of Pragmatic Markers: The Case of <i>You Know</i>	40
ERIK SCHLEEF AND BRADLEY MACKAY	
3 Quotative Variation and Change in French with Additional Insights from Brazilian Portuguese and Italian	61
STEPHEN LEVEY, LAURA KASTRONIC, SALVIO DIGESTO, AND MÉLISSA CHIASSON	
4 Cross-Linguistic Variation in Spoken Discourse Markers: Distribution, Functions, and Domains	83
LIESBETH DEGAND, ZOÉ BROISSON, LUDIVINE CRIBLE, AND KAROLINA GRZECH	
<b>Part II Innovative Variables in English</b>	105
5 An Emerging Pragmatic Marker: Sentence-Final <i>Is All</i>	107
DANIELA KOLBE-HANNA AND LAUREL J. BRINTON	
	v

vi	<i>Contents</i>	
6	“That Is Totally Not My Type of Film”: Innovations in the Intensifier System of UK English KARIN AIJMER	127
7	<i>Uh</i> , What Should We Count? TIM GADANIDIS AND DEREK DENIS	150
8	Modeling Listener Responses MIRJAM ELISABETH EISWIRTH	173
	<b>Part III Language Contact Settings</b>	191
9	<i>You Know</i> in L1 and L2 English CHLOÉ DISKIN-HOLDAWAY	193
10	General Extenders in Bilingual Speech JOSEPH KERN	212
11	The Diverging Paths of Consequence Markers in Canadian French HÉLÈNE BLONDEAU, RAYMOND MOUGEON, AND MIREILLE TREMBLAY	230
12	What Governs Speakers’ Choices of Borrowed vs. Domestic Variants of Discourse-Pragmatic Variables? GISLE ANDERSEN	251
13	A Place for <i>pliis</i> in Finnish: A Discourse-Pragmatic Variation Account of Position ELIZABETH PETERSON, TURO HILTUNEN, AND JOHANNA VAATTOVAARA	272
	Afterword: Future Priorities in Discourse-Pragmatic Variation and Change Research HEIKE PICHLER	293
	<i>References</i>	301
	<i>Index</i>	329

## Figures

---

1.1	Distribution of EP variants in EOE in apparent time	page 23
1.2	Distribution of EP variants by syntactic position in four age groups	25
1.3	Probability of <i>I think</i> (vs. all other EPs) by syntactic position over apparent time	27
1.4	Distribution of EPs by subject complement over apparent time	29
2.1	Significant scales and factors	54
4.1	Frequency and polyfunctionality of DM use in the French dataset	90
4.2	Frequency and polyfunctionality of DM use in the Spanish dataset	93
4.3	Frequency and polyfunctionality of DM use in the English dataset	96
4.4	Frequency and polyfunctionality of DM use in the Polish dataset	98
4.5	Cross-linguistic domain distribution of discourse markers	99
5.1	Distribution of “, is all .” in COCA per million words (pmw) across text types and time periods	114
5.2	Distribution of “that BE all .” in COCA per million words (pmw) across text types and time periods	117
7.1	Proportion of <i>um</i> per speaker, shaded by region	158
7.2	Proportion of <i>um</i> per speaker, shaded by gender	159
7.3	Proportion of <i>um</i> in apparent time	160
7.4	Proportion of <i>um</i> in apparent time, by gender	160
7.5	Proportion of <i>um</i> in apparent time, by position	161
7.6	Proportion of <i>um</i> in apparent time, by position and cliticization	161
7.7	Conditional inference tree for farmers	162
7.8	Conditional inference tree for interviewers	163
7.9	Proportional and normalized frequency of two hypothetical datasets, one where variants are trading frequencies – (a) and (b) – and one where one variant remains static and the other increases – (c) and (d)	165

viii	<i>List of Figures</i>	
7.10	Frequency of <i>uh</i> and <i>um</i> per 1,000 words	167
7.11	Frequency of <i>uh</i> and <i>um</i> per 1,000 words, by position	168
7.12	Frequency of <i>uh</i> and <i>um</i> per 1,000 words, by position and gender	169
8.1	Structural definition and sequential position of LRs	178
9.1	Average rates of <i>you know</i> per 1,000 words across four speaker groups	204
9.2	Frequency of <i>you know</i> by migrants' level of education in the Dublin corpus	204
10.1	Frequency of general extenders in Spanish and English per 10,000 words according to participant, gender, and language dominance	219
10.2	Percentage of general extenders that fulfill a referential function in English and Spanish according to participant, gender, and language dominance	227
11.1	Study design	237
11.2	Variant frequencies in the 1970s and 2010s in Montreal and Welland	239
12.1	Waveform and spectrogram with <i>please</i>	262

# Tables

1.1	Distribution of Earlier Ontario English speakers	page 21
1.2	Mixed-effects logistic regression testing the fixed effects of YEAR OF BIRTH (centered, continuous), POSITION, and their interaction and a random intercept for SPEAKER on the realization of EPs as <i>I think</i>	27
1.3	Mixed-effects logistic regression testing the fixed effects of PERIOD, SUBCOMP, and their interaction and a random intercept for SPEAKER on the realization of EPs as <i>I think</i>	30
1.4	Mixed-effects logistic regression testing the fixed effects of PERIOD, SUBCOMP, and their interaction and a random intercept for SPEAKER on the realization of EPs as <i>I guess</i>	31
1.5	Cross-tabulation of period and grammatical person of complement clause subject in EOE and TEA, showing the percentage of negated EPs relative to all EPs	33
1.6	Summary of the probabilistic niches of <i>I think</i> and <i>I guess</i>	36
2.1	Overview of guises	46
2.2	Guise distribution across eight subsurveys	47
2.3	Twenty-nine scales presented to participants in the survey	48
2.4	Best models for the three factors. Reference levels are neutral guise, Bella, Sick Sam topic, female respondent, middle class, and <i>you know</i> user	52
2.5	Results from the mixed-effects ordinal regression models for scales on which the ratings of <i>you know</i> and neutral guises were judged to be statistically different. References levels are neutral guise, Sick Sam topic, female respondent, middle class, and <i>you know</i> user	53
3.1	Characteristics of the corpora	66
3.2	Variant distribution by age in three varieties of French	71
3.3	Variant distribution by age in Quebec and Acadian French	73
3.4	Two independent multivariate analyses of social and linguistic factors contributing to the selection of <i>être comme</i> in Quebec French and Acadian French (younger speakers only)	75
		ix

x *List of Tables*

3.5	Two independent multivariate analyses of factors contributing to the selection of <i>être comme</i> in Quebec French and Acadian French (younger speakers only)	77
3.6	Variant distributions in Brazilian Portuguese by age	79
3.7	Variant distributions in Italian by age	79
3.8	Summary of major findings in relation to the rate of change exhibited by quotative innovations in French and Brazilian Portuguese	81
4.1	Taxonomy of domains and functions (Crible and Degand 2019b)	88
4.2	Distribution of DMs across domains and functions in the French dataset (dominant functions are highlighted)	89
4.3	Distribution of DM types over discourse functions in the French dataset	90
4.4	Distribution of DMs across domains and functions in the Spanish dataset	91
4.5	Distribution of DM types over discourse functions in the Spanish dataset	92
4.6	Distribution of DMs across domains and functions in the English dataset	94
4.7	Distribution of DM types over discourse functions in the English dataset	95
4.8	Distribution of DMs across domains and functions in the Polish dataset	97
4.9	Distribution of DM types over discourse functions in the Polish dataset	98
4.10	Cross-linguistic distribution of DM functions across discourse domains	101
5.1	Synchronic corpora used in this study	110
5.2	Sentence-final ( <i>that</i> ) <i>BE all</i> in COCA and SOAP (search string: “_vb* all .”)	112
5.3	Sentence-final <i>is/was all</i> in COCA and SOAP (search string “vb* all .”; PM=pragmatic marker)	114
5.4	Co-occurrence of sentence-final <i>is all</i> with <i>just</i>	115
5.5	Co-occurrence of sentence-final <i>that’s/is all</i> with <i>just</i>	118
5.6	Diachronic corpora used in this study	119
5.7	<i>That/this BE all</i> in Late Modern English	122
5.8	Expanded forms with <i>all</i> or <i>what</i> in CEN (free-standing and parenthetical forms)	124
6.1	Number of occurrences of <i>totally</i> used by speakers in each age group in the BNC2014S	133

List of Tables	xi
6.2 The frequencies of <i>totally</i> with different syntactic targets and functions in the BNC1994D and the BNC2014S	134
6.3 The frequency of <i>totally</i> with positive, negative, and neutral adjectives	135
6.4 Occurrences of <i>totally</i> as a modifier of scalar adjectives produced by male and female speakers in different age categories	138
7.1 List of speakers	156
7.2 Cross-community comparison	157
7.3 Organization of the data	166
7.4 Mixed-effects Poisson regression model on non-initial (UHM) counts	169
8.1 Cross-tabulation showing how often each participant takes the speaker and listener role, depending on the gender of their interlocutor	175
8.2 Cross-tabulation of raw token counts of LRs by listener and speaker gender across all interactions	175
8.3 Minimum turn length at which the first LR is predicted based on the zero-model	184
8.4 Predicted number of responses to a 100-word turn based on speaker and listener gender	184
8.5 Fixed effects of the zero-inflated Poisson model predicting how many actions listeners do based on which party has how much impact on the action from a qualitative perspective	186
9.1 Participants in the Dublin and Melbourne corpora	198
9.2 Proficiency levels from the Common European Framework of Reference for Languages (CEFR)	199
9.3 Functions of <i>you know</i> in the Dublin corpus	201
9.4 <i>You know</i> among L1 and L2 speakers	203
9.5 Breakdown of <i>you know</i> by function (House 2009): Dublin corpus	205
9.6 Mixed-effects regression output for <i>you know</i> as illustrator by L1 speakers (Intercept) and L2 speakers in the Dublin corpus	205
9.7 Mixed-effects regression output for <i>you know</i> to refer to shared knowledge by L1 speakers (Intercept) and L2 speakers in the Dublin corpus	206
9.8 Mixed-effects regression output for <i>you know</i> to invite response by L1 speakers (Intercept) and L2 speakers in the Dublin corpus	206
10.1 Frequency of general extenders in Spanish and English per 10,000 words according to participant, gender, and language dominance	218

xii *List of Tables*

10.2	Forms of general extenders in English and Spanish	220
10.3	Length of general extenders in English and Spanish regardless of function	222
10.4	Distribution of referential and non-referential functions of general extenders in English and Spanish according to participant, gender, and language dominance	226
11.1	Corpora	237
11.2	Influence of connector choice according to linguistic function in Montreal 1971 and Welland 1975	240
11.3	Influence of social factors in Montreal 1971	241
11.4	Influence of social factors in Welland 1975	242
11.5	Influence of connector choice according to linguistic function in Montreal 2012 and Welland 2012–2015	244
11.6	Influence of social factors in Montreal 2012	245
11.7	Influence of social factors in Welland 2012–2015	246
12.1	Corpora used in this study	255
12.2	Use of <i>please</i> and equivalent domestic polite expressions in requests (total corpus token frequency in parenthesis)	257
12.3	Use of <i>sorry</i> and equivalent domestic polite expressions in apologies (total corpus token frequency in parenthesis)	264
13.1	Preferred position of <i>pliis</i> and <i>kiitos</i> on a Likert scale of 0 to 3: Can you give me that?	277
13.2	Representative request types in Finnish (non-comprehensive)	278
13.3	Baseline (elicited) positions for <i>kiitos</i> in seven request types	280
13.4	Baseline (elicited) positions for <i>pliis</i> in seven request types	281
13.5	Most frequent request strategies in transit data + lexical politeness markers	284
13.6	Distribution of request types for <i>pliis</i> and <i>kiitos</i> in FinnishWaC	287

## Contributors

---

Karin Aijmer is Professor Emerita at the University of Gothenburg, Sweden. Her research interests focus on pragmatics, discourse analysis, modality, corpus linguistics, and contrastive analysis. Her publications include *Conversational Routines in English: Convention and Creativity* (1996), *English Discourse Particles: Evidence from a Corpus* (2002), and *Understanding Pragmatic Markers: A Variational Pragmatic Analysis* (2013). She is the coeditor of *Pragmatics of Society* (2011) and *A Handbook of Corpus Pragmatics* (2014).

Gisle Andersen is Professor of English linguistics at NHH Norwegian School of Economics in Bergen, Norway. His research interests span topics in pragmatics (discourse markers, interjections, politeness, etc.), language contact, lexicography/terminology, and language for specific purposes. His recent work has focused on ways in which linguistic items that are inherently pragmatic in nature are borrowed across languages. He has studied cases of pragmatic and phraseological borrowing and investigated how the use of borrowed items is constrained by cultural, social, and cognitive factors.

Hélène Blondeau is Professor of French and Francophone Studies and Linguistics at the University of Florida, USA. Her research investigates language variation and change at the individual and community level. She has examined varieties of French in North America with a focus on the Montreal sociolinguistic dynamics.

Laurel J. Brinton is Professor of English Language at the University of British Columbia, Vancouver, Canada. Her research interests include grammaticalization and lexicalization, historical pragmatics, phrasal verbs and composite predicates, verbal aspect, and intensifiers. Her most recent publications include the edited textbook *English Historical Linguistics* (2017) and the monograph *The Evolution of Pragmatic Markers in English: Pathways of Change* (2017) as well as five readers in De Gruyter Mouton *The History of English* series (2017) coedited with Alexander Bergs.

Zoé Broisson was an MA student in Linguistics at the University of Louvain (UCLouvain), Belgium, at the time this volume was compiled. She has since left academia and is now active as a cofounder and linguist in an EdTech startup and UCLouvain spin-off, developing a mobile app that English learners use to boost their pronunciation using voice technology.

Mélissa Chiasson is a project officer in the Office of the Vice-President, Research, at the University of Ottawa, Canada. Her doctoral research targeted sociolinguistic variation and change in Acadian French, as spoken in northeast New Brunswick.

Ludivine Crible earned her PhD in Linguistics at the University of Louvain (UCLouvain), Belgium, in 2017 on the topic of discourse markers and disfluency in spoken English and French. She published several corpus-based papers on the functions and variation of discourse markers, including from a contrastive perspective. After a Marie Curie postdoctoral fellowship at the University of Edinburgh, she is now an assistant professor in French linguistics at Ghent University (Belgium).

Liesbeth Degand is Full Professor of General and Dutch Linguistics at the Catholic University of Louvain (UCLouvain), Belgium. Her research interests are oriented toward the discourse grammar interface, with special attention to the study of discourse markers, discourse segmentation, and clause combining in spoken language. From a methodological point of view, she combines (multilingual) corpus studies with elicitation tasks. She is currently involved in a European project on prediction and alignment in spoken interaction.

Derek Denis is Assistant Professor of Linguistics in the Department of Language Studies at the University of Toronto Mississauga, Canada. His research focuses on the sociolinguistics of innovation and change with special attention to discourse-pragmatic features. His current work investigates the emergence of a multiethnolect in Toronto: Multicultural Toronto English.

Salvio Digesto is a part-time professor in the Department of Linguistics at the University of Ottawa, Canada. His doctoral research addressed variation and change in Italian. He has additional interests in the sociolinguistics of French and English.

Chloé Diskin-Holdaway is Senior Lecturer in Applied Linguistics at the University of Melbourne, Australia, and Chair of the Discourse-Pragmatic Variation and Change (DiPVaC) research network. Her research expertise lies at the interface of sociolinguistics and second language acquisition. She is the (co)author of several articles, including “Going global and sounding local: Quotative

variation and change in L1 and L2 speakers of Irish (Dublin) English” (2019) and “‘Northmen, Southmen, comrades all’? The adoption of discourse like by migrants North and South of the Irish border” (2020).

Mirjam Elisabeth Eiswirth is a postdoctoral researcher at the Sociolinguistics Lab at the University of Duisburg-Essen within the DFG-funded project “Tracing Language Variation and Change across the lifespan in Tyneside English.” Her research focuses on the relationship between interactional structure and language variation at different levels of linguistic architecture, as well as variation and change in the use of interactional strategies. Her work has been published in the *Journal of Pragmatics*, the *Journal of Sociolinguistics*, and with Cambridge University Press.

Tim Gadanidis is a PhD researcher in the Department of Linguistics at the University of Toronto, Canada. His research primarily focuses on language and the Internet, such as how the use and perception of the English discourse-pragmatic markers *um* and *uh* vary across spoken and online (instant-messaging) communication. Most recently, he has used qualitative and computational-linguistic methods to critically analyze discourse about climate change in online communities.

Karolina Grzech specializes in language documentation and description. She received her PhD from SOAS University of London, UK, in 2017 and worked as a postdoctoral researcher at Stockholm University, Sweden. Her main research interests are evidentiality, discourse markers, and analysis of interactive spoken discourse. She conducts fieldwork related to these topics on Quechuan languages in Ecuador. In 2021, she began a postdoctoral project on Quechuan discourse at the University of Valencia.

Turo Hiltunen is University Lecturer in English at the Department of Languages, University of Helsinki, Finland. His research concentrates on corpus linguistics and corpus development, grammar, phraseology, and the analysis of scientific discourse, past and present.

Laura Kastronic is Assistant Professor in the Department of French Studies at the University of Toronto, Canada. Her research interests focus on language variation and change, language contact, bilingualism, and the sociolinguistics of French and English.

Joseph Kern is Assistant Professor in Spanish at the University of Virginia’s College at Wise, USA. His research focuses on

discourse-pragmatic variation in Spanish, especially varieties of Spanish in contact with English and Portuguese, as well as intersections between sociolinguistics and second-language pedagogy. He has analyzed several discourse-pragmatic variables in Spanish, including discourse markers, quotatives, general extenders, and tag questions. He is on the steering committee of the Discourse-Pragmatic Variation and Change (DiPVaC) research network.

Daniela Kolbe-Hanna is a teacher and researcher of English Linguistics at Trier University, Germany, where she received her PhD, in 2008, after an MA in Freiburg in English, Scandinavian and Slavic Studies, during which she studied for a year at the University of Glasgow. Her main research interest lies in the constant variation and change of language. Her publications include the development of the complementizer *as*, variation in complement clauses, and comment clauses in varieties of English around the world.

Stephen Levey is Associate Professor in the Department of Linguistics at the University of Ottawa, Canada. His research interests focus on morpho-syntactic and discourse-pragmatic variation and change, as well as language contact and historical sociolinguistics.

Bradley Mackay is a research assistant in the Department of English and American Studies at the University of Salzburg, Austria. His research focuses on language, gender, and sexuality and he is currently writing his PhD on the sociophonetic variation found among people with divergent sexual identities.

Raymond Mougeon is Professor Emeritus of French Linguistics at York University, Canada. He has carried out extensive research on the evolution of Ontario-spoken French, the ethnolinguistic vitality of Ontario's French-speaking communities and the acquisition of French by Ontario's French immersion students.

Elizabeth Peterson is University Lecturer in the Department of Languages at the University of Helsinki, Finland. She conducts research on language contact and language attitudes and ideologies. She is the author of *Making Sense of "Bad English": An Introduction to Language Attitudes and Ideologies* (2020). She is on the steering committee of the Discourse-Pragmatic Variation and Change (DiPVaC) research network.

Heike Pichler is Senior Lecturer in Variationist Sociolinguistics at Newcastle University, UK. She is the author of *The Structure of Discourse-Pragmatic Variation* (2013), editor of *Discourse-Pragmatic Variation and Change in English* (2016), and founding chair of the international Discourse-Pragmatic Variation and Change (DiPVaC) research network. Her ongoing research focuses

on discourse-pragmatic variation in narrative discourse and computer-mediated communication as well as on language variation and change in old age.

Jan-Ola Östman is Professor Emeritus in Scandinavian Languages at the University of Helsinki, Finland. He is the author and editor of books and articles on pragmatics, discourse, construction grammar, variability, language contact, dialects, socio-onomastics, minority languages, language policy, ideology, and responsibility. He is the coeditor of the John Benjamins book series *Constructional Approaches to Language* and International Pragmatics Association's *Handbook of Pragmatics*. He was President of the International Pragmatics Association in the years 2012–2017.

Erik Schlee is Professor of English Linguistics at the University of Salzburg, Austria. His research focuses on variation and change in dialects of the British Isles, the acquisition of variation, sociolinguistics and perception, and language and gender in educational settings. He is coeditor of *The Routledge Sociolinguistics Reader* (2010) and coauthor of *Doing Sociolinguistics: A Practical Guide to Data Collection and Analysis* (2015).

Mireille Tremblay is Professor of Linguistics at the University of Montreal, Canada. She specializes in language change, morphosyntactic variation, and history of the French language. Her most recent research focuses on the new language practices of Montrealers and documents the changes in progress in the community.

Johanna Vaattovaara is Professor of Finnish language at the Faculty of Information Technology and Communication Sciences at Tampere University, Finland. Her main research areas are linguistic variation and change, language attitudes, and language ideologies. Her PhD on the Tornio Valley (2009) was the first folk linguistic dissertation on the Finnish language. Her more recent work looks at language ideologies in Finnish and on awareness and attitudes toward English loans integrated into Finnish.

## Foreword

### Discourse, Pragmatics and Responsibility

---

In the late 1960s, many scholars within text linguistics in Europe and within generative semantics in the United States became increasingly dissatisfied with a seemingly exclusive focus on the form and structure of sentences in mainstream linguistics. There was a growing need to open up the subject to make it acceptable to talk about context, function and variation in relation to the study of linguistics. Words and sentences have different functions in different contexts.

This turn had already seen the light of day in philosophy (the “later” Wittgenstein, ordinary language philosophy, and finally the influential studies by Searle and Grice); in anthropology (in fact, ever since Malinowski); in translation theory (including Pike’s 1954 [1967] magnum opus); and in ethnography. Even though cultural, social, cognitive, philosophical, and anthropological aspects of language had been actively taken into account in neighboring disciplines and in non-mainstream approaches to linguistics, the stronghold by structuralist and transformational-generative thinking seemed unbeatable: the quest by linguists to be(come) properly scientific had steered the development of linguistic thinking on a route that had taken the very concept of language away from language usage and language users’ adaptability.

Gradually, things started to change in the 1960s: the initial attempt was to build context and function on top of what we already had, that is, adding extra “boxes” or “circles” or “trees” on top of, or around, the structuralist specifications of language as a system, coupled with intricate discussions about compositionality and native speakers’ often very varied intuitions.

Text linguistics developed into discourse analysis, which today includes contextual, political, and multimodal analyses. William Labov (e.g., Labov 1972a) took variation seriously, establishing what we today know as sociolinguistics, and started his and his followers’ search for contextual factors, that is, sociolinguistic variables that could explain linguistic variation: age, gender, race, education, class, among others, to begin with – and new variables are added and tried out constantly. Corpora needed to be gathered and built in order to make it possible for everyone to check everybody else’s data. Generative semantics was for a time declared dead,<sup>1</sup> but it gradually developed into

<sup>1</sup> For elusive, entertaining, and informative accounts of Generative Semantics, see Lakoff (1989), McCawley (1994).

linguistic pragmatics, which in the 1980s increasingly merged with the European, broader approach to language function inspired by Charles W. Morris's three-way distinction between syntactics, semantics, and pragmatics and his view that pragmatics should deal with "all the psychological, biological, and sociological phenomena which occur in the functioning of signs" (1938: 108).<sup>2</sup>

What was soon realized as being more important than adding context and function as viable factors to be referred to – very often "when needed" – was the methodological, epistemological, if not ontological, idea that complementary approaches could be used as tools in the study of language. Once function, variation, and context were taken seriously, linguistic data could be understood and interpreted in very many additional ways. The very concepts of context and function had to be explicated and extended: What *is* context? What *is* function? What *is* variation? What *is* discourse? Context and function for what and for whom – in society, in culture? Further, what are their effects really on language structure and language use? How do language structure and language usage themselves affect context? Some scholars even gave up using the very words *language* and *linguistics* altogether because these had too many connotations relating to studies of explicit word-and-sentence structures of languages. Yet talking about, say, communication, interaction, collaboration, or even Malinowski's (1923) communion, only moved the question in another direction and required understanding and definition of these additional concepts.

In the development of a need to look at function, context and variation as such, pragmatics, sociolinguistics, and discourse analysis have joined forces. Investigations of what goes on implicitly in communication, between the lines of what is actually said as propositional content, have become central. As a result, and in order to tackle such issues, within context-enhanced, interactional, and usage-based approaches to language it has thus become more and more acceptable to see language *from different perspectives*, with each perspective contributing to our overall understanding of, precisely, language.

We consequently find research quests taking semiotic perspectives on language as well as cultural and anthropological perspectives, sociological and philosophical perspectives, cognitive and conceptual perspectives, interactional perspectives, and more. Different perspectives have developed into different models, the results of which have gradually added to our *understanding* of language.

The very concept of "understanding" is important, because ultimately most linguists and students of language working within the humanities and social

<sup>2</sup> This is reflected and reiterated in the mission statement of the International Pragmatics Association to represent "the field of pragmatics, i.e., the **science of language use**, in its widest interdisciplinary sense as a *functional* (i.e., cognitive, social, and cultural) perspective on language and communication" (see <https://pragmatics.international/page/Mission>; emphasis in original).

sciences rarely seek the ultimate truth but, precisely, understanding – understanding of the phenomenon of language; and we have learned that understanding is not easily achieved through deductive reasoning, by setting up a priori models to be applied to instances of communicative expression. Understanding is attained by looking at a phenomenon from different perspectives – ideally at the same time. In practical analyses, however, we take one perspective at a time while simultaneously trying to keep the big picture in mind.

One set of linguistic expressions that very early on challenged the established view that linguistics should first and foremost focus on positive, declarative, transitive, finite clauses with a preset toolbox of word class categories and grammatical functions was the frequently occurring phrases *you know*, *I guess*, *like*, *well*, *oh*, and similar difficult-to-define phrases in terms of predicate logic. Even though school training did its utmost to obliterate these from the vocabularies of young children, the expressions kept occurring – even in adulthood among well-educated people. In fact, these are some of the most interactive features in language, and if linguistics is to deal with ordinary people's language – rather than, or at least in addition to, the written word, focusing on the language of classical authors – these “small words and phrases” had to be taken seriously; and so they have.

A prominent research network that has taken as its task to bring together scholars and promote research in this field is the Discourse-Pragmatic Variation and Change (DiPVaC) network, with a specific focus on, precisely, discourse-pragmatic features like the discourse markers or pragmatic particles already mentioned, which have as their main function to express interpersonal, (con)textual, and affective functions in communication. The present volume brings together an exciting collection of chapters by members and collaborators of this network.

Yet the road traveled to where we are at the moment has not always been an easy and navigable one. Robin Lakoff's (1972) article was an early attempt at getting scholars' attention to turn to context and language function, and since it was published in the journal *Language* it was not completely marginalizable, albeit that it was published at a time when language function and the very field of pragmatics were often scorned by mainstream linguists.

Lakoff's article may have opened the eyes of the English-language community of linguists to what she called pragmatic particles, but there had indeed been much earlier studies of (pragmatic) particles, in other languages, especially in (literary versions of) earlier forms of languages, and in languages that had not been previously described, where particles and clitics were a permanent challenge to the anthropologically minded linguist (see, e.g., Denniston 1934 on Greek particles; Kinkade 1976 on Interior Salishan particles). German particles received attention very early on – probably because many of the German particles also functioned as conjuncts, and the task was to explicate

the differences in meaning and function of these linguistic expressions (cf. Arndt 1960; Schubiger 1965; Kriwonossow 1966). In fact, in the 1960s and 1970s, there was constant debate about whether the particles were grammatical and syntactically specifiable (as modal particles, as attitudinal adverbs, as conjunctive or connecting particles, or, for example, in the case of Japanese, as subject or topic particles) or whether their function was better described as more pragmatic.<sup>3</sup>

Even though the study of conversation analysis had started to make an appearance as early as in the mid-1970s (with Sacks et al. 1974 being the major article that reached the general linguistics audience – also published in *Language*), and although there were MA theses and PhD theses written on particles in different languages, a more intensive and specific research focus on pragmatic particles as expressions with different interactional functions did not properly start until the early 1980s. This was also the time when scholarship in pragmatics as a field of study on a par with phonology, morphology, syntax, lexicology, and semantics was beginning to become established – and with it not only the International Pragmatics Association but also book series in the field, notably the John Benjamins series *Pragmatics and Beyond*, as part of which I published one of the first studies that attempted to describe the intricacies of the pragmatic particle *you know* (Östman 1981): its use in interactions, in narratives, and in child language acquisition as well as differences in use with respect to gender.

It was important at the time to try to come up with a definition of the pragmatic particle I was investigating.<sup>4</sup> Thus, I suggested the following definition as the general, prototypical meaning of *you know*:

The speaker strives towards getting the addressee to cooperate and/or to accept the propositional content of his[her] utterance as mutual background knowledge.

(Östman 1981: 17)

The definition is a product of its time (e.g., with regard to gender), but one of the most crucial indications to support and validate this definition was the data for “the same story” being told twice by the same teller to the same addressee. In the first telling of the story, that is, when the story is new to the addressee, the teller uses an abundance of *you knows*, striving to check the level of understanding of the addressee, pleading for cooperation, and seeking to get the addressee to take for granted the tenability of what they are being told. In the second telling, however, when the teller knows that the addressee knows the story, the teller uses hardly any *you knows*.

<sup>3</sup> See Weydt’s influential 1968 book on *Abtönungspartikel* (Ger. *Abtönung*, “shading, toning down”).

<sup>4</sup> This is, of course, still important, but notions like particles being multifunctional and polysemous are much more readily acceptable nowadays.

Conversation analysts in their very detailed ethnomethodological analyses of everyday interaction, focusing on wanting to answer the question of “Why that now?” with respect to the occurrence of a linguistic item,<sup>5</sup> necessarily also paid attention to pragmatic particles as well as to silence, supportive elements, filled pauses, and the like – all of which are also at the center of interest for the DiPVaC research network. When John Gumperz, on the basis of his experiences in India and of language contacts in London, wanted to make sociolinguistics become more “discoursy,” a new branch called interactional sociolinguistics came into being; and his new book series at the time, *Studies in Interactional Sociolinguistics*, published what became the standard reference book for studies in this field for a long time, Deborah Schiffrin’s (1987) *Discourse Markers*. The term was in accordance with Gumperz’s (1984) book *Discourse Strategies*, and the very term “discourse marker” has today become as frequent – if not dominant – as that of “pragmatic particle.”<sup>6</sup>

The rest, as one says, is history.<sup>7</sup> These early studies effectively opened up a field of inquiry that is still extremely vibrant forty years later. This edited volume and the DiPVaC network build on these earlier and subsequent studies and are, in that sense, one of the follow-up outcomes of the early pioneers’ work.

It is difficult to make any predictions about the future of this line of work. One thing is certain, though, and that is that, despite the multitude of books and articles on pragmatic particles, discourse markers, discourse-pragmatic features, and the variability and adaptability of human communication, there is still a lot to do. That is of course another good reason for having a network like the DiPVaC, which, according to their mission statement,<sup>8</sup> “provides a platform for the dissemination and discussion of new research findings, the formation of new research collaborations, and the promotion of the field within and beyond linguistics.”

All scholars working in this field would agree that pragmatic particles/discourse markers do not partake in establishing or affecting the propositional content of expressions they are used in connection with; and there are, of course, other characteristics: they are usually short, they are inflexible (which is the definition of “particle”), and they modify, qualify (etc.) something else.

<sup>5</sup> This was in stark contrast to the transformation-generative grammarian’s “Have you made a generalization today?”.

<sup>6</sup> Schourup (1983) – another early study – used the term discourse particle. Other terms include IFIDs (illocutionary-force indicating device), discourse connectives, speech-act adverbials, interjections, and pragmatic markers.

<sup>7</sup> For an overview of publications, see Aijmer and Simon-Vandenberg (2009), but any respectable journal dealing with language function publishes articles on discourse markers, and references in the chapters in this volume clearly also include more recent contributions.

<sup>8</sup> See the DiPVaC research network website: [www.dipvac.org](http://www.dipvac.org).

Various contextual factors and variables have laudably been suggested as influencing their function.

Are there more general characterizations we can make, though? In my own work (cf. Östman 1982, 1986, 1995), I have talked about pragmatic particles as having an implicit anchoring function, as being the implicit anchors of messages par excellence – they are the markers that let us see what goes on “between the lines” of what we say, the verbalized connections to the implicit communication we engage in whenever we speak: our connection to the background society and culture we take for granted, to the systems of politeness and turn-taking we utilize, to the involvement, affect and emotions we might not want to verbalize directly.

I would like to conclude with a few words about the last of the trinity in the subheading to my Foreword, viz. *responsibility* (cf. Popper 1987; Solin and Östman 2016). I want to do this to suggest one direction for future research on discourse-pragmatic features.

If we take the general function of pragmatic particles to implicitly anchor a message, and to communicate information about (1) cultural coherence, (2) interaction-politeness, and (3) affect-involvement, then they are the surface manifestations par excellence of the underlying, implicit aspects of messages. The definition of pragmatic particles that I have taken as a guideline in my own research runs as follows:

Pragmatic particles are (verbal) elements in language that have as their primary function to implicitly anchor utterances vis-à-vis the communicative restraints of a culture and society, the demands of aspects of interactive politeness, and the prevalent types of affect and involvement.

To this, I offer the addendum that access to and employment of pragmatic particles are part of our competence – as is pragmatic information generally. That is, every expression that takes linguistic form is a crystallization of some structurally, cognitively and/or socially relevant factor (for human expression); and what joins these three perspectives together is precisely the understanding of how communicators handle responsibility in communication.

Indeed, I feel the expression – or rather, negotiation – of responsibility is perhaps the most salient common aspect of the function of pragmatic particles as communicating implicit, subconscious messages: you implicitly express what level and kind of responsibility you want to take, give, have, and so on with pragmatic particles.

For instance, and very briefly, if we take another look at the abovementioned narrative data and my definition of the function of *you know* from the point of view of responsibility, we could say that, in the first telling of the story (with an

abundance of *you knows*), the responsibility of what actually took place in the story was solely the teller's; but, in the second telling, when the addressee also knew the story, the responsibility for the story was shared by the teller and the addressee.

In other words, the focus on responsibility as a tool would perhaps not add a completely new interpretation, but it does add an additional *perspective*; and having access to a multitude of perspectives is crucial in order to get “the full picture” – especially so in discourse-pragmatic studies.

I think the gist of doing discourse-pragmatic work is precisely to keep focusing on what has traditionally been talked about as the periphery, the margins, the fringes. Context is also a perspective on language – it is not something “outside of” language. This insight alone will give us a deeper understanding of the function(s) of language.

JAN-OLA ÖSTMAN

## Acknowledgments

---

The editors of this volume are indebted to the contributing authors, whose expertise and dedication to creating high-quality research remained a chief motivator throughout the editorial process. Our sincere thanks also go to the reviewers of the chapters in this volume as well as the external reviewers of the book proposal and complete manuscript. Their care and attention to their task lifted the writing and reporting of our research to a different level. We thank our editor at Cambridge University Press, Helen Barton, for her enthusiasm and encouragement from the initial stages through to the end of the publication process. We also thank our editorial consultant, Isabel Collins, for tirelessly, effectively and promptly responding to our many questions, as well as the editing and Integra production team for their careful work during the final stages of the project. Most of the chapters in this volume emerged from presentations at the DiPVaC4 conference held in Helsinki, Finland, in 2018. For this event, we thank the University of Helsinki, the university's Faculty of Arts and the Federation of Finnish Learned Societies for their support. We likewise thank the audience and participants at DiPVaC4 for their comments and feedback on our work. Finally, our thanks to Heike Pichler for sharing her expertise and knowledge on several occasions throughout the crafting of this volume. Her work and example as a leader in the field continue to inspire us.

## Abbreviations

---

AmE	American English
AP	adjectival phrase
AusE	Australian English
BLV	Belleville
BrE	British English
CA	conversation analysis
CEFR	Common European Framework of Reference for Languages
CI	credible interval
CMC	computer-mediated communication
CP	complementizer phrase
CRE	Constant Rate Effect
DiPVaC	Discourse-Pragmatic Variation and Change [network]
DM	discourse marker
DPM	discourse-pragmatic marker
ELF	English as a Lingua Franca
EON	Eastern Ontario
EP	epistemic parenthetical
F-INT	female interviewer
IrE	Irish English
L1	first language
L2	second language
LOESS	locally estimated scatterplot smoothing
LR	listener response
M-INT	male interviewer
MC	middle class
NIA	Niagara
NNS	non-native speakers
NP	noun phrase
NS	native speakers
PAF	principal axis factoring
PM	pragmatic marker (Chapter 5)
PM	politeness marker (Chapter 13)
pmw	per million words
PP	prepositional phrase
RL	recipient language
SES	socioeconomic status
SL	source language

xxviii     *List of Abbreviations*

SLA	second language acquisition
TCU	turn construction unit
TRP	transition-relevance point
UMC	upper middle class
WC	working class