

Contents

Preface	vii
Editors	viii
Notes on Contributors	ix

- 1 Corpus Linguistics as Digital Scholarship: Big Data, Rich Data and Uncharted Data 1
Terttu Nevalainen, Carla Suhr and Irma Taavitsainen

PART 1

Evidence from “Big Data”

- 2 Big Data: Opportunities and Challenges for English Corpus Linguistics 29
Antoinette Renouf
- 3 Corpus-based Studies of Lexical and Semantic Variation: The Importance of Both Corpus Size and Corpus Design 66
Mark Davies
- 4 Empirically Charting the Success of Prescriptivism: Some Case Studies of Nineteenth-century English 88
Lieselotte Anderwald
- 5 *Warn Against -ing*: Exceptions to Bach’s Generalization in Four Varieties of English 109
Mark Kaunisto and Juhani Rudanko

PART 2

Evidence from “Rich Data”?

- 6 Commonplace Books: Charting and Enriching Complex Data 133
Thomas Kohnen
- 7 Mining Big Data: A Philologist’s Perspective 151
Tanja Rütten

- 8 Function-to-form Mapping in Corpora: Historical Corpus Pragmatics and the Study of Stance Expressions 169
Daniela Landert

- 9 Scholastic Argumentation in Early English Medical Writing and Its Afterlife: New Corpus Evidence 191
Irma Taavitsainen and Gerold Schneider

PART 3

Evidence from Uncharted Data and Rethinking Old Data?

- 10 Language Surrounding Poverty in Early Modern England: A Corpus-based Investigation of How People Living in the Seventeenth Century Perceived the Criminalised Poor 225
Tony McEnery and Helen Baker
- 11 An Information-Theoretic Approach to Modeling Diachronic Change in Scientific English 258
Stefania Degaetano-Ortlieb, Hannah Kermes, Ashraf Khamis and Elke Teich
- 12 Academic Vocabulary in Wikipedia Articles: Frequency and Dispersion in Uneven Datasets 282
Turo Hiltunen and Jukka Tyrkkö
- 13 *Words (don't come easy):* The Automatic Retrieval and Analysis of Popular Song Lyrics 307
David Brett and Antonio Pinna
- 14 Charting New Sources of ELF Data: A Multi-genre Corpus Approach 326
Mikko Laitinen, Magnus Levin and Alexander Lakaw
- Index 351