large-scale corpora, and those that work with complex statistical evaluations, and of course those that combine any of these aspects. In sum, I believe this is a great volume rightfully dedicated to David Denison, 'a master at this integrative approach to the study of language change and the history of English' (Mair, p. 361).

Reviewer's address:
Osnabrück University
Institut für Anglistik und Amerikanistik
Neuer Graben 40
D-49069 Osnabrück
Germany
abergs@uni-osnabrueck.de

References

Anderson, Stephen & Paul Kiparsky (eds.). 1973. *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston.

Brinton, Laurel. 2017. *The evolution of pragmatic markers in English: Syntactic origins and pragmatic development*. Cambridge: Cambridge University Press.

Corbett, Greville & Michael Noonan (eds.). 2008. *Case and grammatical relations: Studies in honor of Bernard Comrie*. Amsterdam: John Benjamins.

Darnell, Regna, James D. McCawley & Arnold Zwicky (eds.). 1992 [1971]. *Studies out in left field: Defamatory essays presented to James D. McCawley on the occasion of his 33rd or 34th birthday*. Amsterdam: John Benjamins.

Denison, David. 2006. Category change and gradience in the determiner system. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 279–304. Oxford: Blackwell.

Hickey, Raymond & Stanislav Puppel (eds.). 1984. *Language history and linguistic modelling: A Festschrift for Jacek Fisiak*. Berlin: de Gruyter.

Huddleston, Rodney & Geoffrey K. Pullum *et al*. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Radford, Andrew. 2004. *Minimalist syntax: Exploring the structure of English*. Cambridge: Cambridge University Press.

(Received 20 October 2020)

Axel Bohmann, *Variation in English worldwide: Registers and global varieties* (Studies in English Language). Cambridge: Cambridge University Press, 2020. Pp. xvii + 250. ISBN 9781108751339.

Reviewed by Melanie Röthlisberger ⓘ, University of Zurich

Bohmann's empirical study analyses the role of variety and register in structuring variation in World Englishes from a usage-based perspective. The number and types of

linguistic features compared, and the amount of varieties and registers examined, are so far unprecedented in research in World Englishes (WE) and register studies. Lying at the intersection of WE research, corpus-based register analysis and aggregation-based analysis of linguistic variation (dialectometry), this study challenges the importance of geography in WE research and emphasizes the importance of register.

Chapter 1 ('Introduction', pp. 1–8) outlines how the study connects three strands of research on the theoretical and methodological level, namely WE research, corpus-based register analysis (à la Biber 1988) and dialectometry. In that way, the multi-feature approaches of dialectometry and corpus-based register studies can benefit research in WE to define varieties not in terms of nation-state bounded entities but by taking a bottom-up perspective that relates varieties in terms of linguistic similarity. While such quantitative dialectological methods have been applied to study the two largest standard varieties (American and British English), Bohmann claims that little is known about global patterns of (register and geographic) variation – a lack that his study aims to overcome. The last part of the chapter is dedicated to the research objectives and a short outline of the book chapters.

In chapter 2 ('The world of English: Variation in geography and register', pp. 9–40), Bohmann summarizes theoretical accounts of WE research and register studies, and justifies the inclusion of online Twitter data in his study. First, Bohmann critically summarizes the main tenets of three influential theoretical models of WE, Kachru's Circle Model, Schneider's Dynamic Model and Mair's World System of Englishes. While the usefulness of these models has been repeatedly empirically tested, most WE studies have done so by limiting their focus to a single or a small set of linguistic features and by mainly ignoring the importance of the situational–functional context of language use. In contrast, corpus-based register studies take register as the fundamental determinant of variation. Work in that area compares the frequencies of a large set of linguistic features across different texts in order to determine the underlying dimensions of register types, thereby often ignoring the importance of geographic stratification. By adding the register dimension to his geographic analysis, Bohmann aims to explore systematically variability in the linguistic features that instantiate registers across space. The chapter finishes by justifying the inclusion of Twitter data in the analysis with practical reasons and theoretical considerations.

In chapter 3 ('Quantifying linguistic variation', pp. 41–58), Bohmann describes the methods that linguists have applied to quantify linguistic variation, the theoretical paradigms they stem from, and their merits and limitations, before deciding on one method, namely multidimensional (MD) analysis. While methods from other paradigms (variationist sociolinguistics, corpus-linguistic analysis and dialectometry) provide a similar rigorous approach to analyse variation, the methods are deemed insufficient for Bohmann's purpose in their limited focus on one register, a comparatively small set of linguistic features, decontextualized language usage, reliance on survey data instead of naturalistic data, or a geographically limited set of varieties. Traditional register studies, on the other hand, include register as an

important determinant of variation with multidimensional scaling techniques – a method that can provide both the aggregate perspective and linguistic particulars.

Chapter 4 ('The space of variation in the present study', pp. 59–100) describes the corpora used, the 236 linguistic features selected and how the methodological parameters of the factor analysis were set, and provides a first overview of the ten dimensions derived from the MD analysis. The ten selected corpora from the *International Corpus of English* (ICE) series sample a broad range of different registers and national standard varieties, namely Canada, Great Britain, Hong Kong, India, Ireland, Jamaica, New Zealand, Philippines, Singapore and the USA (only written). The ICE corpora's focus on standard varieties and their potential lack in accurately representing the sociolinguistic reality of a country is compensated for by adding a large corpus of geolocated Twitter messages. Linguistic features were selected on the basis of existing literature in register analysis, dialectometry and World Englishes. These include grammatical and morphosyntactic features, elements of discourse organization and of word formation, discourse properties, and stance-taking devices, always also including non-standard spelling conventions. The last section of the chapter is devoted to a comprehensive explanation of the technical details of exploratory factor analysis with a particular focus on the number of factor dimensions, the factor extraction method, the method of factor rotation and the factor scoring method. The ten resulting dimensions, determined on the basis of the full dataset, are labelled by Bohmann as INVOLVED VS INFORMATIONAL PRODUCTION, COLLABORATIVE COMMUNICATIVE ORIENTATION, CONCEPTUAL VS CONCRETE INFORMATIONAL FOCUS, CANONICAL NARRATIVE FOCUS, SITUATIONAL ANCHORING OF REFERENCE, COLLOQUIAL MARKEDNESS, EXPLICIT STANCE-MARKING, FUTURE-ORIENTED DISCOURSE, ASSERTION OF FACTUAL VALIDITY and ADDRESSEE-ORIENTATION. Chapters 5, 6 and 7 are then all devoted to the discussion of these ten dimensions.

Chapter 5 ('General situational dimensions of variation', pp. 101–24) discusses the first three dimensions. These reflect general situational properties that tend to characterize prototypical written and prototypical spoken texts. Each dimension is introduced by first summarizing its salient linguistic features, next by exploring the distribution of dimension scores by modality, text type (according to the ICE corpus structure) and variety, and finally by providing sample texts that highlight the use of the linguistic features in context. The first dimension (INVOLVED VS INFORMATIONAL PRODUCTION) is characterized by a preference for linguistic features that relate to nominal style, to personal detachment and to structural complexity, with Inner Circle varieties preferring a more involved style compared to Outer Circle varieties. The second dimension (COLLABORATIVE COMMUNICATIVE ORIENTATION) measures communicative co-presence/ co-production. This dimension differentiates between (graphemic and conceptually) spoken text types and (graphemic and conceptually) written text types. As such, scripted monologues seem to contain more properties of written communication while written correspondence has more discourse properties of oral communication. On the geographic level, spoken IndE seems to be less collaborative and communicative oriented and spoken CanE more so. The third dimension (CONCEPTUAL

VS CONCRETE INFORMATIONAL FOCUS) differentiates between abstract-conceptual and concrete-referential informational focus. Salient linguistic features serve to emphasize and elaborate on basic information with the help of adjectival, adverbial and (prepositional) phrasal modification. Variation between modalities and varieties seems to be inconsistent on the whole, but a closer look reveals that differences between spoken and written registers are more pronounced in Inner Circle varieties than in Outer Circle varieties.

The three dimensions discussed in chapter 6 ('Register-specific dimensions', pp. 125–46) reflect highly specific discourse types – often reflective of one register sampled. The first of these dimensions, CANONICAL NARRATIVE FOCUS, is characterized by linguistic features that relate to the narration of past events, such as third-person pronouns and verbs in the past (perfect). It also includes standardized spelling, which is reflective of the editorial process when publishing literary works. Literary texts (novels and short stories) thus score highest on this dimension, with the other registers and especially tweets having comparatively lower scores. Regional variation seems to reflect a variety's developmental stage according to Schneider's model: phase 5 varieties (Inner Circle) and phase 4 varieties (e.g. India, Jamaica, Singapore) tend to have a higher narrative score across all modalities than phase 3 varieties. Bohmann explains this difference with the emergence of a native literary tradition in phase 4. The next dimension, dimension 5 (SITUATIONAL ANCHORING OF REFERENCE), designates discourse that refers to a temporal or spatial environment. Salient linguistic features include preposition sequences, place and time adverbials, which represent strategies to define (spatial or temporal) physical relations. These features seem to be characteristic of spontaneous monologues, particularly of spontaneous sports commentaries which contribute the most to the register's high score. Regional variation is again explained by varieties' developmental stage: varieties in phase 5 score highest in this dimension and varieties in phase 3 have the lowest dimension scores because this dimension's specific discourse style – reflective of linguistic diversification – can only have emerged by phase 5 in Schneider's model. The sixth dimension (COLLOQUIAL MARKEDNESS) reflects colloquial and conversational style. The dimension comprises salient linguistic features that represent interpersonal and informal aspects of discourse, some of which conform to non-standard norms and constitute stylized features (e.g. first-person pronouns, non-standard second-person plural pronouns, or contractions). Twitter scores highest on this dimension and is highly distinct from other registers. Spontaneous conversations, fiction writing and correspondence also score comparatively highly due to their interpersonal discourse characteristics. From the regional perspective, North American varieties (US, Canada, Jamaica) tend to exhibit higher scores on this dimension, which Bohmann explains with increased colloquialization in American English and the status of American English as a linguistic hub in Mair's System of World Englishes.

Chapter 7 ('Dimensions with other patterns of distribution', pp. 147–72) discusses the remaining dimensions, which reveal diverse patterns of distribution. Dimension 7 (EXPLICIT STANCE-MARKING) is marked by analytic and explicit stance-taking linguistic

features, which is characteristic of face-to-face conversations and loose information-packaging. Public dialogues and persuasive writing score highest on this dimension. Geographic differences set Inner Circle varieties apart from Outer Circle varieties, with Inner Circle varieties scoring higher across all modalities, particularly in the spoken register. This is also reflective of the varieties' developmental stage: stage 5 varieties have the highest scores and stage 3 varieties the lowest. Dimension 8 (FUTURE-ORIENTED DISCOURSE) is the least conclusive dimension. Registers that score highly on this dimension tend to include some type of addressee-oriented discourse but the connection to future-oriented discourse is unclear. This dimension seems to capture geographic variation, with Inner Circle varieties scoring below 0 and Outer Circle varieties scoring above 0. This regional difference is most discernible in private dialogues. Dimension 9 (ASSERTION OF FACTUAL VALIDITY) seems to reflect argumentative discourse. Salient linguistic features include the use of BE as a main verb, demonstrative pronouns and definite articles, and adverbials indicating epistemic certainty. Student writings and public dialogues score the highest on this dimension and the Twitter data the lowest. Bohmann explains the former with the argumentative nature of student essays and public dialogues and the latter with the more emotional rather than fact-oriented argumentation abounding on Twitter. The five Outer Circle varieties score highest on this dimension, particularly so in the spoken register. The final dimension (ADDRESSEE-ORIENTATION) includes linguistic features that are indicative of an addressee-oriented communicative style, such as second-person pronouns, conditional subordinators *if* and *unless*, and modals *can* and *may.* According to Bohmann, these features can be used as polite hedging devices in addressee-oriented communication. High scores on this dimension are reached by instructional and fictional writing – two registers that include explicit directions to readers/addressees and where the audience is well defined. No discernable and clear-cut regional differences can be observed in this dimension.

Chapter 8 ('Discussion: Feature space and geographical space', pp. 173–90) compares the study's findings to earlier register studies and WE work. Bohmann observes that the study's MD analysis largely overlaps with previous results, validating the current findings. Among other things, the comparison highlights the fact that some dimensions capture register-specific discourse while others capture a geographic signal in the data (e.g. FUTURE-MARKING DISCOURSE). Bohmann then examines the ways in which the ten English varieties relate to each other along the extracted register dimensions and discusses how this relationship can contribute to theoretical models of World Englishes. The linguistic relationship between varieties is explored using a hierarchical cluster analysis on the mean scores of all features by variety. The emerging dendrogram separates the US and British English(-influenced) varieties (British, Irish and New Zealand English) from all other varieties. The three remaining clusters group Indian and Canadian English, Singapore and Hong Kong English, and Jamaican and Philippines English together. These clusters are supported only to a certain extent by the Circle and the Dynamic models. Due to the clustering's very broad perspective on intervarietal differences, Bohmann also zooms in on differences in spoken, written and

Twitter data. This analysis highlights the usefulness of Kachru's Circle and Schneider's Dynamic Model for accounting for intervarietal differences in the three modalities, while the System of World Englishes supports findings for the dimension COLLOQUIAL MARKEDNESS (with American English as a hub). Bohmann ends the chapter by arguing that register seems to be a much better predictor of variation than geography, as shown with the test statistic that calculates the relationship between the dimension score and variety/geography. He concludes that any aggregate labels such as 'Hong Kong English' fail to take intravarietal differences in situational contexts into account and that register should play a more prominent role as a vehicle of linguistic variation in World Englishes research.

Finally, chapter 9 ('Conclusion', pp. 191–6) provides a short summary of the study's results and stresses again the exploratory nature of the study. On a theoretical level, the Dynamic Model turned out to be the most useful model for accounting for the observed patterns of variation due to its explicit reference to the sociohistorical reality of varieties in different developmental stages. This model has shown that, overall, the varieties in early developmental phases tend more towards formal and informational linguistic patterns while phase 5 (or Inner Circle) varieties exhibit a more affective and involved style. This difference highlights the fact that the situational context should feature much more prominently in WE research.

Bohmann's study provides a much-needed in-depth analysis of variation across geographic and situational–functional space in World Englishes. The study succeeds in highlighting intravariety heterogeneity and thus challenges traditional WE research that views variety as a homogeneous entity. In that respect, Bohmann's study follows a recent trend in linguistics that aims to combine the register perspective with the WE perspective, but none of the previous work takes as many features, as many different text types (including online data) and as many varieties into account. The study's focus is thereby mostly on situational contexts and the discourse setting, while variety seems to play a negligible role in accounting for variation in the feature space (only prominent in two dimensions). This also means that regional differences between registers and text types and, for instance, the extent to which the same registers are comparable across World Englishes are only marginally discussed. This is mostly due to the methodological approach chosen (MD analysis) but also due to the aggregate perspective that the study takes, a limitation that Bohmann himself points out. Bohmann addresses this latter issue – the aggregate perspective – by providing additional detailed discussion of texts that exemplify the extracted dimension with a close look at the dimension's characteristic linguistic features. This detailed exemplification adds a discourse-analytic aspect to the analysis that is sometimes missing in quantitative WE research. The aggregate perspective necessarily also conceals the extent to which a different set of linguistic features might have characterized dimensions extracted separately by variety (instead of from the full dataset of all ten varieties). And it also conceals the comparability of the registers sampled, that is, the extent to which the ICE data (sampling started in the 1990s) are

comparable to the Twitter data (sampled in the 2010s), an issue that Bohmann does not address.

World Englishes research can clearly benefit from this work as it highlights the importance of situational context when comparing varieties of English and the effect that the choice of register as representing 'a variety' can have in WE theorizing. While Bohmann clearly states the benefits of this kind of multi-register approach to WE studies, the analysis presented is clearly also beneficial for register studies that, similarly, often view a text type as a homogeneous entity rather than as the heterogeneous register it could be from a WE perspective. In addition, Bohmann's study could also be useful for dialectometric research, a linguistic paradigm that Bohmann also saw as intersecting with his work (see chapter 1). For instance, dialectometry could be potentially inspired by the range and types of linguistic features analysed (dialectometric research often does not concern itself with discourse-marking devices) and the bottom-up approach taken by Bohmann (using the linguistic features to determine the space of variation).

Finally, Bohmann's study sets the pace for future systematic quantitative research in World Englishes that aims to compare varieties on linguistic grounds sampling from naturalistic language data (for a similar attempt but focusing on probabilities rather than frequencies and on a small set of linguistic variables see Heller 2018; Röthlisberger 2018; Szmrecsanyi *et al.* 2019). What Bohmann's work and other quantitative research in World Englishes have in common is that the theoretical models of World Englishes that they draw on often fail as a perfect explanans for their findings. Rather, the quantified linguistic distance between varieties is more complex than such models suggest, often depending on the linguistic features analysed, the measures used for comparison (frequencies or probabilities), the varieties and registers examined, and the representativeness of the corpus texts collected. Most studies observe a dichotomy between L1 and L2 varieties (or Inner vs Outer Circle) (e.g. Szmrecsanyi *et al.* 2019); sometimes we can also observe a North American cluster or detect the influence of British English on its former colonies (as in the case of Bohmann's study with British and New Zealand English and the North American varieties clustering together). Other variety-groupings are harder to interpret and explain on sociohistorical grounds. As Bohmann himself concedes (p. 194): 'The most important finding is that patterns of individual dimensions may differ and require specific explanations, rather than being easily subsumed under one static image of WE relations.'

Reviewer's address:
University of Zurich
Deutsches Seminar
Schönberggasse 2
CH-8001 Zurich
Switzerland
melanie.roethlisberger@uzh.ch

References

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

Heller, Benedikt. 2018. Stability and fluidity in syntactic variation world-wide: The genitive alternation across varieties of English. PhD dissertation, KU Leuven.

Röthlisberger, Melanie. 2018. Regional variation in probabilistic grammars: A multifactorial study of the English dative alternation. PhD dissertation, KU Leuven.

Szmrecsanyi, Benedikt, Jason Grafmiller & Laura Rosseel. 2019. Variation-based distance and similarity modeling: A case study in World Englishes. *Frontiers in Artificial Intelligence* 2. DOI: 10.3389/frai.2019.00023

Isabel Moskowich, **Begoña Crespo**, **Luis Puente-Castelo** and **Leida Maria Monaco** (eds.), *Writing history in late Modern English: Explorations of the Coruña Corpus*. Amsterdam: John Benjamins, 2019. Pp. vii+278. ISBN 9789027204240.

Reviewed by Nicholas Groom, University of Birmingham

This book heralds the release of the *Corpus of History English Texts* (CHET). CHET is a recent addition to the *Coruña Corpus of English Scientific Writing*, an ongoing series of diachronic corpora compiled by the MuStE research group at the University of A Coruña (www.udc.es/grupos/muste). Each corpus in the Coruña series covers the same historical period of 1700 to 1900, and has been carefully designed so as to be directly comparable with other corpora in the series. While its name may suggest an exclusive focus on STEM disciplines, the Coruña Corpus project actually aims to provide resources for studying the evolution of academic writing in English across the whole disciplinary spectrum, from Astronomy to Philosophy, and from Life Sciences to History, the subject of the current volume.

As with the two previous edited collections emerging from this project (also published by John Benjamins), the aim of the book is not to provide a coordinated set of claims about the development of historical writing during the period covered by CHET, but rather to showcase the variety of research questions that can be addressed using the corpus, either by itself or in combination with other corpora in the Coruña stable. Ultimately, the book is successful in this aim, although it must be said that this is only after a somewhat shaky start.

The Foreword (by Javier Pérez-Guerra) informs the reader that the opening chapter will be by MuStE research group lead Isabel Moskowich, and that this chapter will provide 'some background about the discipline in the period, the trigger for compilation [sic] and [a] description of the corpus material' (p. 2). However, chapter 1 in the published volume is not by Isabel Moskowich, and does not actually do any of the abovementioned things. Instead, chapter 1 presents the reader with a sprawling and at