
Reviewed by Warren Maguire, University of Edinburgh

This book analyzes the relationships between a range of dialects of British English using an innovative approach—the dialectometrical analysis of morphosyntactic variation in a corpus of recorded speech. Prior to Szmrecsanyi’s research, dialectometry—the simultaneous mathematical analysis of many geographically different linguistic varieties across a wide range of linguistic features—had largely been directed at lexical and phonetic variation (see Maguire & McMahon 2011, though see Spruit 2006 for a notable exception). S’s research is the first large-scale dialectometrical investigation into morphosyntactic variation and, to add to its novelty and rigor, it uses as its database an extensive corpus of naturalistic conversational speech. As such, it is a very welcome addition to our understanding of the relationships among British English dialects, to geographical morphosyntactic variation, and to our understanding of (the limits of) dialectometrical and quantitative linguistic methods more generally.

The book is divided into nine chapters. Ch. 1 introduces the rationale for the study and surveys previous analyses of relations among English dialects, both traditional analyses based on isoglosses and dialectometrical analyses based on lexis and phonetics. Ch. 2 describes the database underlying S’s study, the Freiburg English dialect project and corpus (FRED; see Kortmann
& Wagner 2005) and the basic principles and methods of corpus-based dialectometry. Ch. 3 describes the fifty-seven morphosyntactic features investigated, showing their geographical patterning and comparing this to previous accounts of these features. Ch. 4 goes beyond the analysis of individual features and examines aggregate morphosyntactic similarities and differences among all thirty-four dialects, and between them and Standard British and Standard American English. Ch. 5 investigates whether the relationships between varieties are gradient, and the degree to which the linguistic differences between them correlate with geographical distance. Ch. 6 takes this analysis further, exploring the extent to which dialect clusters are apparent in the data, using hierarchical agglomerative cluster analysis. Ch. 7 seeks to determine whether there are any correlations between the distributions of the different features, using principle component analysis. Chs. 8 and 9 summarize the findings of S’s analysis, including an assessment of the significance of the geographical dimension in morphosyntactic variation. These are followed by three appendices, giving summary statistics and—a major feature of the book—a series of twenty-nine color maps and four dendrograms illustrating the results of the analysis.

Grammatical variation in British English dialects is a clear and extremely informative account of dialectometrical methods. Chs. 2 and 4 in particular are an excellent introduction to dialectometry to anyone new to the field, while Chs. 5, 6, and 7 provide a very readable survey of advanced dialectometrical techniques, though the mathematical details of multidimensional scaling, hierarchical agglomerative cluster analysis, and principle component analysis are skirted over, which may be a good or bad thing, depending on your perspective. These chapters are really brought to life by the full color maps and dendrograms in Appendix C, which beautifully and clearly illustrate the results of the techniques used by S. In addition to the many grayscale figures in the rest of the book, they form a mini linguistic atlas of S’s data. The book is worth it for these aspects alone, and I would certainly recommend it to students as essential reading for an understanding of dialectometry. Ch. 3 is also a useful summary of key patterns of morphosyntactic variation in British English dialects, though, given the large number of features described, it cannot go into a great deal of depth; however, S very helpfully provides plenty of references to previous research for each of the features for the interested reader to pursue further.

Despite the many excellent aspects of this book, there are issues with the analysis and interpretation of the data, which means that the conclusions drawn by S raise as many questions as they answer. First, although S includes fifty-seven morphosyntactic features in his analysis in order to determine linguistic similarities and differences among the dialects under investigation, he finds (41–69) that only fourteen of these have a significant geographic distribution in his data (i.e. the other forty-three show no significant geographic patterning), even though most of them have been found by previous researchers to vary in geographically interesting ways (which is why S included them in his analysis in the first place). The reason for this disparity is not explained, and the effect of only fourteen of the features showing significant geographic patterning is not fully explored. S makes the valid point (70) that ‘[w]e would, however, like to emphasize at this time that although individual features, when analyzed in isolation, may not have a regionally significant frequency distribution, they may still be implicated in regionally distinctive feature bundling’ (see also the discussion in Maguire & McMahon 2011:98), but the problem still remains that S’s analysis of the FRED data gives rather different results than we might expect, and readers are left to determine for themselves why this is the case. Second, the linguistic relationships of some locations, especially (but not only) those in northern England, are decidedly odd. For example, on p. 85 S notes that the varieties that are closest to Standard British English are ‘the relatively young dialects in the Scottish Highlands and on the Hebrides’ (a not-unexpected result) and ‘the measuring points in the North of England’ (Westmorland, Yorkshire, and Lancashire are in the top six most similar dialects, morphosyntactically, to Standard British English—see p. 86 and Map C13 in Appendix C). This is surprising, given that these dialects are often quite strikingly divergent from southern British English (see, for example, Shorrock 1999). Why this might be so is not explained at all by S, but an exploration of such patterns is surely necessary when they appear to contradict what we know about linguistic relations between varieties of English from other studies.
The two issues with S’s study outlined above relate to two aspects of his analysis that should have been addressed more fully in the book. First, it is difficult to know how to assess S’s results. One issue that other researchers conducting dialectometrical analyses of phonetic data have pursued is how we can tell whether the results of their analyses actually mean anything. Heeringa and colleagues (2002:446) suggest that the results of dialectometrical analysis should be compared to a ‘gold standard’, which they describe as ‘a classification of language varieties with which (nearly) all experts agree’. Such a gold standard might include uncontroversial classifications arrived at by a wide range of more traditional studies. Although traditional classifications of British English dialects based only on morphosyntactic features are lacking, the fact that S’s overall results do not match previous classifications of British English dialects based on other criteria particularly closely and that many of S’s features fail to pattern in the same way as they do in other studies suggests that the results of his analysis are not likely to meet such a gold standard. The second issue is that S does not spend enough time in his methods section discussing exactly how common a feature must be at each data point in his corpus for it to give robust results (though see pp. 24–27 for some related discussion). Table A.1 (Appendix A) reveals that some of the features are very common indeed (e.g. feature 14, total text frequency of forms of to be, with 90,779 occurrences in the corpus as a whole), while others are very rare indeed (e.g. feature 7, synthetic adjective comparison, with 109 occurrences in the corpus as a whole). Of the fifty-seven features listed in Table A.1, nineteen have frequencies across the whole corpus of less than 500, while a further thirteen have frequencies of less than 1,000. These numbers might not seem particularly low, but when we divide them among thirty-four dialects, we are, in many cases, dealing with very low numbers indeed. How many occurrences of each feature are needed before their frequencies are no longer subject to the vagaries of text length and chance? The absence of a discussion of this issue in the book is unfortunate, especially since token frequency is a recurrent problem in corpus-based morphosyntactic analysis (see Buchstaller et al. 2013), and one wonders whether some of the unexpected results (the lack of geographical patterning of many features and the unusual linguistic relations of some dialects) might be the result of essentially random patterning of low-frequency features.

These two issues raise questions as to the efficacy of the kind of corpus-based dialectometry approach adopted by S. Nevertheless, this is an important book in the growing field of dialectometry. It is a clear and engaging demonstration of some key dialectometrical techniques that allow us to probe the relations between varieties of a language, and between those and geography, in new ways. The maps in Appendix C are a highlight. How the issues identified with S’s study might be addressed will determine the future of the corpus-based approach to morphosyntactic variation he advocates. Regardless, this book is undoubtedly a major contribution to the field and will be a catalyst for further research.

REFERENCES


Reviewed by Tor A. Åfarli, Norwegian University of Science and Technology NTNU*

This book is the second title in the series ‘Key topics in syntax’, which, according to the text facing the title page of the book, aims to ‘bridge the gap between textbooks and primary literature’ and ‘consists of accessible yet challenging accounts of the most important issues, concepts, and phenomena to consider when examining the syntactic structure of language’. The key topic treated in this book is clause structure. This topic is viewed from the point of view of contemporary generative syntax, which means the minimalist program, broadly understood. Thus, this book traces the developmental paths of minimalist syntax, highlighting the various theoretical shifts and problems, and the empirical data that led to them. Sometimes, substantial research findings are pointed out and described, but quite often no well-founded theoretical conclusions are reached and the issue at hand remains unresolved, albeit still illuminated.

In the preface, the author mentions (x) that the book ‘is more of a textbook than a monograph’, and it is stated (xi) that the intended audiences for the book are advanced students and colleagues. Also, it is mentioned that a familiarity with the basics of generative and minimalist syntax is presupposed. The book focuses on clause structure in English, other languages being considered whenever it is appropriate, which happens quite frequently. In addition, all of the central chapters contain a substantial section titled ‘Cross-linguistic observations’, which broadens the empirical perspective further. Even though the theoretical focus clearly is on contemporary minimalist theory, facts and analyses from the history of generative syntax are often considered as well, which is welcome, since it adds depth to the discussion.

The book contains seven chapters. Each chapter (except Ch. 7, the conclusion) has a section called ‘Discussion points’ and suggestions for further reading. The introductory chapter (Ch. 1) provides a compact introduction to generative syntax, including a short history of the role of universal grammar in generative syntax, the change from rule-based to principles-and-parameters-based syntax, and phrase structure theory from phrase structure rules, via X-bar principles, to features and bare phrase structure. There are brief discussions on topics like VP-shells, Agree, probe-goal checking, Merge, minimalism vs. cartography, the linear correspondence axiom, phases, and the nature of features, to mention some of the topics that are covered. Chs. 2–6 comprise the core of the book. Ch. 2 discusses some of the basics concerning the clause and presents the three structural layers of the clause, viz. the VP-layer, the TP-layer, and the CP-layer, which provide the organizing arrangement for the chapters that follow. Thus, Chs. 3–5 are devoted to each of the three layers, starting with the VP-layer, and Ch. 6 is concerned with the connection between the layers.

Ch. 2 starts out with an examination of the main clause, focusing on main clauses in English. The basic characteristics of the main clause in English are said to be that it has a certain mood and

* Thanks to Terje Lohndal for useful comments to an earlier version of this review.