

## 120 *Reviews*

Taavitsainen, Irma, Merja Kytö, Claudia Claridge and Jeremy Smith (eds.). 2015. *Developments in English: Expanding Electronic Evidence*. Cambridge: Cambridge University Press. ISBN 978-1-107-03850-9.

The twofold aim of this volume is (i) to shed light on how the empirical, increasingly corpus-based approach “has transformed our understanding of both contemporary and historical varieties” and (ii) “to provide an up-to-date, forward-looking account of what is taking place in corpus-based research into the English language” (p. 1). Given the considerable impact that corpus-linguistic methods have had on English linguistics, a volume that achieves this aim would be a very welcome contribution to the field. The contributions to the book are organized into four parts with three or four articles per part; each part also begins with a brief introduction by one of the editors. In my review, I shall address each contribution in turn before concluding with an overall evaluation of the volume.

Before Part I, an introduction by the editors sets the scene by describing the rise of corpus linguistics within English linguistics, outlining problems as well as possibilities that have resulted from this sea change in the context of recent trends in the field, and explaining on what basis the contributions to the volume have been organized into parts. The introduction is worth reading not only as a valuable gateway to the volume itself, but also for the lucid way in which it outlines the growth and modern challenges of corpus-based research.

Part I, “Linguistic Directions and Crossroads: Mapping the Routes”, begins with an introduction by Merja Kytö, in which the three contributions to Part I are claimed to “illustrate the complexity of such a primary issue in corpus linguistics as how to approach one’s data” (p. 13). Charles F. Meyer’s “Corpus-based and Corpus-driven Approaches to Linguistic Analysis: One and the Same?” is the first contribution. Meyer’s chapter explores whether there is really a divide between corpus-based approaches (CBAs) and corpus-driven approaches (CDAs). Meyer links CBAs to deduction, since the formulation of one or several hypotheses precedes the corpus analysis, while CDAs are inductive in that “theoretical claims ... are postulated only after an intensive corpus analysis has been conducted” (p. 15). Criticism of CBAs typically focusses on the danger that theories, hypotheses, and grammatical categories (including those underlying corpus annotation)

which are formulated prior to the examination of the data will result in restricted and/or biased analyses. Based on an examination of two case studies, Meyer argues that even supposedly corpus-driven analyses are likely to be based on some theoretical framework and that such a framework “need not limit or inhibit a corpus analysis” (p. 19). Meyer also devotes special attention to the annotation of corpora by means of tagging and/or parsing, a practice to which proponents of CDAs are often opposed because it is seen as forcing “the analyst into a particular paradigm of grammatical analysis” (p. 24). Meyer acknowledges that annotation is potentially problematic, especially when the tagger or parser excludes relevant linguistic tokens from retrieval, but he also notes that it is possible to ignore annotation when it is not useful. As regards both underlying theoretical frameworks and corpus annotation, Meyer’s position is thus that they do not introduce bias into analyses because the analyst is free to discard or modify pre-existing frameworks. Meyer concludes that CBAs and CDAs have so much in common that these terms could be abandoned in favour of a common term like *corpus approaches*, that the debate between their proponents “is partially ideological”, and that it would be preferable to evaluate corpus studies “based on their individual merit” (p. 28). Meyer’s point that CDAs typically also contain a theoretical framework such as word classes is valuable, and his examination of CBAs and CDAs is insightful and of clear importance to the development of the field.

Stefan Th. Gries’s chapter “Quantitative Corpus Approaches to Linguistic Analysis: Seven or Eight Levels of Resolution and the Lessons They Teach Us” deals with a fundamental question in corpus linguistics, viz. what “the role of quantitative/statistical work and overall methodological sophistication” (p. 29) is in this field of scholarship. Gries’s discussion results in eight “lessons” that can help scholars to improve their analyses. Some of these—e.g. Lesson 1, “[f]requencies should always be augmented by, or checked against, dispersion measures” (p. 34)—will be familiar to many researchers (which does not of course mean that they are unnecessary; frequencies are often presented without dispersion measures in linguistics). Others, such as Lesson 6, “[d]o not lump together contexts of (co-)occurrence but distinguish them and their type–token distributions and consider their dispersion/uncertainty” (p. 41), require considerable statistical

sophistication on the part of the researcher; while the lesson itself looks possible to grasp, the discussion of *measures of uncertainty* and *(relative) entropy* that precedes it is more challenging. Indeed, parts of the statistical discussion in Gries's contribution remain opaque to me; to the extent that I am representative of the target group of the volume, it might have been preferable to focus on fewer "lessons" and treat those more extensively, with more background information provided. However, readers with the necessary statistical knowledge will find a wealth of very valuable suggestions for how to improve their analyses in this contribution.

As pointed out by the editors, the final contribution to Part I, "Profiling the English Verb Phrase over Time: Modal Patterns" by Bas Aarts, Sean Wallis, and Jill Bowie, differs from the two preceding chapters in being a case study rather than a discussion of methodology. The study is based on the parsed *Diachronic Corpus of Present-day Spoken English*, which includes material from the *London–Lund Corpus* (1960s and 1970s) and the British component of the *International Corpus of English*; the authors also compare their dataset with results from previous work on the *LOB*, *FLOB*, *Brown*, and *Frown* corpora. The data were extracted using the *International Corpus of English Corpus Utility Program (ICECUP)*, which allows for powerful and precise grammatical searches. The results show that spoken BrE and written BrE appear to be more clearly distinct than written BrE and written AmE; for instance, core modals are more frequent in the spoken material. The authors also chart trajectories of individual modals relative to the overall set of modals. This method reveals some interesting differences compared with frequencies related to the number of words; for example, while "*can* appears static" in terms of its normalized frequencies, its proportion of all core modals increases in all three diachronic comparisons. As the authors point out (pp. 56–57), even such figures cannot tell us whether specific modal uses are replacing others, as modal auxiliaries are not semantically equivalent; however, the results strongly suggest, for instance, that *will* is replacing *shall*. Finally, modal patterns in the two spoken subcorpora are explored in great detail: the *ICECUP* software enables the authors to retrieve different patterns such as "modal + main verb" and "modal + subject + X + VP" automatically. This fine-grained analysis reveals a number of diachronic and synchronic patterns; for instance, the simple declarative pattern "modal + main verb" displays

the most dramatic decrease. The authors' detailed analysis is especially valuable since the object of investigation is spoken English, which may be "forecasting changes in written British English" (p. 75). The authors successfully demonstrate the value of parsed corpora—and of advanced software that can retrieve information from them—to diachronic analyses. Taken together, the three chapters in Part I make a valuable contribution to corpus-linguistic methodology.

Claudia Claridge introduces Part II, "Changing Patterns". Even though the contributions to Part II are quite diverse, they are claimed to have common denominators in that they address "shifts in semantic-pragmatic and grammatical function" and "point to the importance of language contact in the evolution of the forms under discussion" and of "paying attention to functional and stylistic varieties of the language" (p. 80). The first chapter in Part II is Minoji Akimoto's "On the Functional Change of *Desire* in Relation to *Hope* and *Wish*", which covers the period from Middle to Present-day English and is based on the *Helsinki Corpus*, *ARCHER*, *FLOB*, and *Wordbanks Online*. The first item under discussion is the parenthetical use of these verbs, as in *He will, I hope, come today*. Akimoto shows that *desire* and *wish* parentheticals "are always accompanied by *as*" and that *hope* "has developed its parenthetical function since the Early Modern English period" (p. 86). Akimoto then turns to the competition between *to*-infinitive and *that*-clause constructions for these three verbs; among other things, he links the fact that no parenthetical marker with *desire* has developed to the disappearance of *that*-clause complementation with a zero subordinator with this verb (e.g. ... *he desired*  $\emptyset$  *he would please to sit down and take share on't*). Akimoto also suggests that the preference for *desire* over *hope* and *wish* in *have* composite predicate constructions (e.g. *have a desire to*) is linked to the higher proportion of nominal uses of *desire* and to the French origin of that verb. Akimoto provides a careful account of the trajectories of these three verbs, richly illustrated with corpus examples. However, I would have liked to see more discussion of other competing expressions in this semantic field, as their distribution may affect all three verbs under discussion; Akimoto notes that "the functional changes of these verbs are by no means local" (p. 82) and mentions the frequency of the verb *want* as a probable reason for the decline of *desire* + *to*-infinitive, but one wonders about other expressions as well, such as volitional *would*.

In “From Medieval to Modern: On the Development of the Adverbial Connective *Considering (That)*”, Matti Rissanen charts the development of *considering (that)* “from the Middle English borrowed verb *consider* to the grammaticalized subordinator and preposition”. Rissanen makes use of the Invited Inferencing Theory of Semantic Change (IITSC), which “presupposes active roles from both speaker/writer and addressee/reader” (p. 98), and draws on several historical corpora for data. Owing to the short time span between the first occurrence of *consider* as a verb and of *considering* as a connective, Rissanen argues that the grammaticalization of the connective started in Anglo-Norman rather than Middle English, although the use of the verb *consider* may still have affected the further development of the connective. Rissanen shows that early forms of *consider* occur primarily in official texts; however, the relatively high frequency of the verb in letters is interesting in that “the cognitive aspect of discourse” in such texts may tally with the occurrence of “cognition-based connectives” (p. 107). Rissanen’s trajectory of development for *consider* into a connective goes via non-personalized uses of the verb where “the individual personal subject as the ‘performer of the cognitive act of considering’ has disappeared or lost its significance” (p. 109). In Early Modern English, both the verb and the connective become more frequent and less restricted in terms of genre. However, *considering* then decreases in frequency during the Modern English period, which may be due to competition with other expressions such as *regarding*. Rissanen suggests that the low frequencies of the connective “in comparison with the other uses of the verb *consider* ... may explain why *considering* was never completely grammaticalized” (p. 114). Rissanen’s thorough discussion of the development of the connective is convincing, and the combination of IITSC and corpus linguistics seems very promising. However, a more in-depth analysis of individual corpus examples than what is provided in his chapter would arguably be a desideratum for such a combined analysis. As Rissanen notes (p. 99), theoretically speaking, studies based on IITSC require “suggestions concerning both the speaker’s intended implications and the inferences created in the addressee’s mind”. Combining such an approach with corpus linguistics would require detailed discussion of key tokens where addressees’ “inferences may result in new interpretations and innovative uses” (p.98).

Manfred Markus's chapter "Spoken Features of Interjections in English Dialect (Based on Joseph Wright's *English Dialect Dictionary*)" (*EDD*) concludes Part II. Markus argues that Wright's "overall output takes us as close to the spoken language as we can possibly get" (p. 117) in Late Modern English studies and that the use of interjections in past speech may be different from that in Present-day English. Markus starts his account with a discussion that problematizes the issue of how to define interjections. The ensuing quantitative survey of the *EDD Online* shows that the full text of the dictionary contains a little more than a thousand interjections, 737 of which are headwords marked as belonging to that word class. The most frequent initial letters of interjections in the *EDD Online* are *h*, *w*, and *s*. Markus attributes the high frequency of interjections in *h* to the articulation of /h/ as part of "strong breathing (inhaling or exhaling), which is a spontaneous reaction in emotive situations" (p. 123) and argues that this also explains why interjections in *w* are frequent: some of them begin with *wh*, which would have been pronounced with initial /hw/ in several varieties. The popularity of interjections in *s* is suggested to be due to hissing playing a role in "our basic inventory of words" (p. 124). While these explanations seem plausible, it would have been interesting to see corresponding figures for initial letters of all words in the dictionary; for instance, *s* is a common first letter of words in English in general, and the high frequency of interjections beginning with *s* might partly be attributable to that fact. A look at the first 26 interjections alphabetically indicates that "they do not adhere to any particular pattern of word formation", although "the main morphological quality of interjections may be seen in their readiness to profit from conversion and the merging of phrases and even clauses" (p. 126). However, I miss a systematic study of a random subsample of the entries rather than an alphabetical list, since interjections beginning with the same letter may constitute a biased sample to generalize from. A division into simple primary interjections (e.g. *ay*), simple secondary interjections (e.g. *allow*), which originate in other parts of speech, and complex interjections (e.g. *adone*), which incorporate phrasal or clausal material, reveals that nearly a third of the interjections are words from other parts of speech that have got a secondary function as interjections in dialectal usage. Markus also examines the semantics and pragmatics of interjections, e.g. the

overlap between interjections and exclamations and “the twofold role of interjections, expressive/emotive and appellative/phatic” (p. 131).

Part III, “Pragmatics and Discourse”, is introduced by Irma Taavitsainen, who notes that these fields are “newer areas of concern for corpus linguists than morphology, syntax, or semantics” and that “[c]orpus pragmatics is in a dynamic stage of development” (p. 137). The first chapter in Part III is Laurel J. Brinton’s “Interjection-based Delocutive Verbs in the History of English”. Brinton thus treats “verbs converted from interjections with the meaning ‘to say or utter [interjection]’” (p. 140), e.g. *to hey*. After introducing and discussing the category of delocutive verbs, Brinton introduces the Delocutive Light Verb Construction (DLVC), which “consists of a verb of saying in combination with a noun, pronoun, particle, or interjection” (p. 145), e.g. *to say hello*. DLVCs are a possible source of delocutive verbs. Brinton’s survey of interjection-based delocutive verbs is based on a large number of corpora; the list of search terms was based on verbs listed as stemming from interjections in the *OED* and on previous research. While there appear to be no relevant items in Old English, Middle English evinces both delocutive verbs such as *fie* and DLVCs like *crien alas*. A number of interjections, e.g. *hush*, also begin to be used as delocutive verbs in Early Modern English. Brinton notes that her corpus data seldom allowed her “to antedate the *OED* examples”, which she attributes to delocutive verbs being “highly salient” and therefore “likely to catch the attention of the dictionary citation collectors” (p. 150). In Late Modern English, the number of interjection-based delocutive verbs increases dramatically, although Brinton notes that some of this increase may be due to the greater diversity and number of textual witnesses for this period. In contrast, relatively few new types arise in Present-day English. Despite the fact that DLVCs are frequent in Middle English, the period before delocutive verbs begin to expand, Brinton argues that the former are an unlikely source of the latter and that delocutive verbs arise independently either through conversion from interjections or as a back-formation from the *-ing* form, given that several delocutive verbs appear for the first time in this form. Brinton also discusses, among other things, the importance of the distinction between the actional sense ‘to perform the action(s) implied by “X”’ and the “saying” sense ‘to say “X”’ of delocutive verbs to the issue of lexicalization (only verbs that have lost the “saying” sense are argued to

be fully lexicalized). Brinton's paper is a lucid and highly readable account of a class of verbs that, although their token frequency is very low, is "represented by a wide variety of types" (p. 140).

Andreas H. Jucker's chapter "*Uh* and *Um* as Planners in the *Corpus of Historical American English*" (*COHA*) aims at investigating the usage of these two planners and comparing it with their distribution in speech. Jucker provides an initial account of speaker evaluation and awareness of planners, which also covers psycholinguistic research into why they occur. Then follows an account of the various functions planners may fulfil in speech. As *uh* and *um* have other uses in addition to their function as planners, subsets of each decade covered by *COHA* were gone through to provide estimated frequencies of the two planners in the corpus; the results indicate that the frequency of the planners rises steeply in the twentieth and early twenty-first centuries (*uh* is ahead of *um* in this development, even though *um* is virtually the only type represented before 1900), but also that they remain far more frequent in corpora of spoken English. Jucker attributes this to the negative stereotypes associated with these planners, which can only sparingly be used by authors for characterization, "presumably because of their foregrounding function and their disproportionate salience in written language" (p. 177). The most important functions of *uh* and *um* in *COHA* are hesitation and planning; the planning process may concern anything from lexical items to entire utterances, in which case planners often occur in connection with awkward or untrue utterances. Jucker's well-informed study provides an appealing mix of quantitative information and close readings of corpus examples.

Thomas Kohnen's "Religious Discourse and the History of English" concludes Part III. As Kohnen notes, "an analysis of religious discourse offers the unique possibility of tracing linguistic phenomena in one domain, covering the whole history of the English language" (p. 178). Despite this centrality, the religious domain is not well represented in historical corpora, something the *Corpus of English Religious Prose* (*COERP*) helps to rectify. Kohnen divides the discourse world of Christianity into three spheres: God addressing the Christian community, the Christian community addressing God, and members of the Christian community addressing one another. Differences and similarities among genres can be explained partly in terms of which sphere they belong to. Genres in *COERP* are grouped into spheres and

classified according to whether they are core genres, which are “a central part of the religious domain and common religious practice” (e.g. sermons), minor genres, which “are used only in special institutions or by specially qualified people” (e.g. monastic rules), and associated genres, which originally existed outside the religious domain but have come to be associated with it (e.g. pamphlets) (p. 180). Kohnen’s survey of existing research based on *COERP* addresses the various text functions of religious prose, interactive features such as directive speech acts, and the continuum of genres that can be established with regard to how linguistically conservative religious prose is. Kohnen then presents a new study of the associated genres of letter pamphlets and prefaces compared with the core genre of religious treatises in sixteenth- and seventeenth-century English; the study is based on parts of *COERP* and on secular treatises from the *Helsinki Corpus*. Kohnen focusses on three features: address terms (second-person pronouns and nominal address like *brother*) and stance marking (first-person constructions such as *I think, I profess*, etc.) “reflect the involved nature of the genres”, while third-person *-th* vs. *-s* is “diagnostic of religious texts” (p. 184). The frequency of address terms rises over time in prefaces, which Kohnen attributes to “a growing focus on the addressee” (p. 185); differences are also attested between the associated genres and the core genre of religious treatises. As regards first-person stance marking, all genres except religious treatises display marked increases in frequency, which tallies with results of previous research; the core genre thus comes across as conservative but does not seem to influence the associated genres. Finally, while religious treatises are quite advanced with regard to the incoming *-s* form early on, they do not generalize this form in the late seventeenth century, unlike the associated genres—especially letter pamphlets—and secular treatises. Kohnen attributes the differences between core and associated genres to the more secular and associated/involved nature of the latter. Kohnen’s convincing discussion is admirably connected to both previous linguistic research and the historical setting of the texts examined; one would, however, have liked to see raw frequencies of the features on which he focusses, given the relatively small text samples (Kohnen mentions on p. 188 that 4,640 address terms occurred in his material, but otherwise appears to report normalized frequencies and percentages only).

The very valuable resource that is *COERP* also raises a question of general concern for diachronic corpus linguistics. As mentioned above, Kohnen makes the important point that religious discourse provides researchers with an opportunity to study material from the same domain across a large time span, which means that the material is highly comparable over time. But the other central desideratum in corpus compilation, viz. representativeness, also needs to be addressed. The question arises how representative religious discourse is of the total textual universe of which it is a part, and whether the representativeness of religious texts changes over time; Kohnen also addresses such issues in his discussion of the relationship between a core religious genre and the emerging standard.

The fourth and last part of the volume, “World Englishes”, is introduced by Jeremy Smith, who notes that the contributions to this part focus on Late Modern English. Smith attributes the recent increase in scholarly attention paid to Late Modern English to developments in corpus linguistics that have facilitated analysis of large quantities of data and to “decentering” trends in colonial and postcolonial studies. Susan Fitzmaurice’s “History, Social Meaning, and Identity in the Spoken English of Postcolonial White Zimbabweans” is the first contribution to the section. Fitzmaurice begins by discussing the different identities associated with the terms “Zimbabwean” and, in particular, “Rhodesian” from a historical perspective. She then turns to the issue of whether these identities can be differentiated phonetically. Starting out from a Zimbabwean informant’s comments on differences between “Zimbo” and “Rhodie” accents with regard to the KIT, DRESS, TRAP, BATH, and NURSE vowels, Fitzmaurice is interested in how Zimbabwean informants see the speech of their peer group “as part of a recognizable register” and in “the extent to which they adapted their discourse and lexical register to accommodate their interviewer” (p. 211); however, the latter question is not explicitly addressed in this study. Her informant pool is made up of 19 Zimbabweans born between 1979 and 1986, who were interviewed about their background and asked to read a passage and a wordlist; most of the interviews were carried out by two informants originally interviewed by Fitzmaurice. Extracts from the interviews illustrate how the informants negotiate their identity as white Zimbabweans, many of whom are now living elsewhere. Fitzmaurice’s account of this issue is fascinating and

illuminating, although it arguably fits the volume less well than the other contributions examined so far. Parts I–IV of the volume are said to deal “with various aspects of diachronic corpus linguistics” (p. 7); but Fitzmaurice’s material is not explicitly diachronic, and while it may be stored electronically, there appears to be nothing in her methodology that requires a corpus-linguistic perspective. (This, however, does of course not detract from the value of her study as such.)

In “Singapore Weblogs: Between Speech and Writing”, Andrea Sand argues for “a context-oriented approach” to the study of New Englishes (p. 224). Sand focusses on the variable of text type; the particular text type examined is “weblogs posted by Singaporean speakers of English”, which was chosen since computer-mediated communication “may promote the use of non-standard varieties in writing” (p. 225). Sand outlines the linguistic landscape of Singapore, with special attention to the use of English and of Singlish, a “creoloid” or “semi-pidgin” that “shows signs of massive restructuring due to the influence of Malay and the various varieties of Chinese spoken in Singapore” (p. 227). Sand’s data come from the Singapore component of the *International Corpus of English (ICE-SIN)* and from the *Corpus of Singapore Weblogs (CSW)*. The three features analysed are discourse particles borrowed from languages other than English spoken in Singapore, zero constituents (copula, subject, or object), and *like* as a discourse marker, hedge, or quotative; the last feature is argued to be characteristic of informal English overall. While discourse particles (e.g. *lah* and *lor*) are attested in *CSW*, the spoken component of *ICE-SIN* contains larger numbers of types as well as tokens. As regards the analysis of zero constituents, the absence of overt subjects and objects that are recoverable from context is a feature of Malay and Chinese, while copula deletion is characteristic of several English-based creoles. A total of 40,000 words from *CSW* and from conversation in *ICE-SIN* were analysed manually to retrieve such zero constituents. There were tokens in both corpora, but again, the frequency was higher in conversation. Zero subjects were the most frequent type, which Sand attributes to them often being easily recoverable. However, Sand measures these features in terms of their normalized frequency per 10,000 words: the overall frequency of subjects, copulas, and objects will thus influence her results. A better method would have been to calculate the percentage of subjects, copulas, and objects that lacked

surface realizations; Sand (p. 234) notes that such an analysis is forthcoming. As regards zero objects, it is also possible that Sand's examples might occasionally allow other explanations; for instance, in *It's about spirits and ghosts kind of thing You watch Øuh*, the question arises whether *You watch* may be a relative clause postmodifying *kind of thing*, in which case Standard English would license a zero relative marker (it is possible that more context would disambiguate the example). As for *like*, finally, both the hedge/emphasis function and the quotative function are represented in conversation as well as weblogs, though they are more frequent in the former. Sand takes the consistent result that all features examined were (i) attested in both conversation and weblogs but (ii) less common in the latter text type to "support the claim that CMC could indeed be one of the factors promoting Singlish in written usage" (p. 236), which seems eminently reasonable.

The third contribution to Part IV is Raymond Hickey's "Mergers, Losses, and the Spread of English". Hickey focusses on mergers "either of vowels or of consonants which were distinguished in the varieties of English taken to overseas locations during the colonial period" (p. 237); however, the FOOT–STRUT split is also treated. His chapter aims at addressing the mergers, the phonological motivation behind them, and their role in the phonology of the varieties that have them. Hickey first treats non-conditioned mergers, which are "characterized by the fact that the two segments which become identical have not done so due to their phonotactic environment" (p. 238). The three mergers taken up are WHICH–WITCH, HORSE–HOARSE, and MEAT–MEET. Hickey demonstrates convincingly that the first merger is best seen as a shift from /hw/ to /w/ in WHICH. As regards HORSE–HOARSE, which concerns /ɔ:/ vs. /o:/ merging as /o:/, Hickey argues that the development may be internally motivated in that it is a counterpart of the MEAT–MEET merger among front vowels: varieties with the HORSE–HOARSE merger would "have a more symmetrical distribution of vowels across phonological space" (p. 241). The next section addresses mergers "before /r/ or its reflex in non-rhotic varieties" (p. 242) such as POUR–POOR and MERRY–MARRY–MARY. The pre-nasal PEN–PIN merger is seen as based on auditory factors: "nasal raising of /e/ to /i/ ... can be seen as a kind of assimilation maximizing the distance between the first and second formants in anticipation of the distance between the two with nasals" (p. 246). Pre-lateral mergers involving loss of length distinctions

such as FEEL–FILL are also addressed. Hickey explains the lack of a HORSE–HOARSE distinction in overseas varieties except some conservative American English ones in terms of when the areas were first settled: only North America would have been settled early enough for the distinction to have applied to contemporary British varieties. The presence of the FOOT–STRUT split in every L1 variety of English outside the British Isles despite the rarity of lowering of /ʊ/ to /ʌ/ in languages generally is explained in terms of the salience of the absence of the split to speakers with the split, whose accent would have dominated colonial administration. Hickey’s proposed reasons for the mergers all seem convincing.

The last contribution to Part IV—and the volume—is William A. Kretzschmar, Jr.’s “Complex Systems in the History of American English”, which “discusses implications for the initial formation of American English and its varieties as the product of random interactions in a complex system between speakers of different input varieties of British English” (p. 251). Kretzschmar argues that, as the population of early English settlements in North America was quite varied in its linguistic make-up and subject to continuous change owing to high mortality rates and continued emigration, a complex system of speech interactions resulted. Because English was the majority language of the settlers, who largely replaced the indigenous population, a variety of English became the everyday language of every colony. Since speech is a complex system, Kretzschmar argues, the variants of any linguistic feature would then be distributed “according to the nonlinear A-curve pattern simultaneously at every level of scale”; different sets of variants would be top-ranked in different localities, and “a particular set of variants emerged as top-ranked elements at the highest level of scale, American English” (p. 257). The complex system would also allow input from non-English sources. In addition, Kretzschmar relates his account to Edgar Schneider’s “Dynamic Model” of the evolution of new varieties and discusses, among other things, the spread of patterns inland from the east coast during westward migration, the retention of particular linguistic features of settlement populations, continued change in twentieth-century American English, and variation in present-day communities as predicted by complexity science. Kretzschmar’s chapter is an

important contribution to theory formation regarding the establishment of extraterritorial varieties of English generally.

Taken together, the standard of the contributions to *Developments in English* is impressive. In addition to the value of the empirical research presented, several contributions open up new vistas into English diachronic linguistics from a theoretical and/or methodological perspective. The quality of the editing is also very high; infelicities are rare and, mostly, trivial. One of the few cases where a reader may be misled is on p. 44, where Stefan Th. Gries discusses  $\beta$ -persistence:

a structure  $x$  increases the probability of a similar structure  $y$  at the next point where  $x$  competes with a functionally similar structure.

Unless I misunderstand the definition, the second “ $x$ ” should be  $y$ .

My only real criticism concerns the place of Part IV in the volume as it is described by the editors. It is arguable that the contributions to Part IV do not address diachronic corpus linguistics, despite the editors’ statement (p. 7) that all sections deal with one or several aspects of this topic. It has already been mentioned that Fitzmaurice’s contribution is not clearly diachronic or corpus-based; in addition, the corpus-linguistic perspective is not apparent in Hickey’s (clearly diachronic) contribution, and Sand’s corpus-based paper does not focus on diachronic data. (Kretzschmar’s contribution, while not explicitly corpus-based, uses material such as the *Linguistic Atlas Project* to make a diachronic point.) This should not be taken as criticism of the studies themselves, however; the overall scholarly value of these contributions is considerable, and they clearly add to the appeal of the volume as a whole. The editors have succeeded in gathering many facets of linguistic research together into a volume which, while it is essential to scholars within English historical corpus linguistics, will be of great value to research on the English language in general.

*Erik Smitterberg*  
Uppsala University