

**Bas Aarts, Joanne Close, Geoffrey Leech and Sean Wallis** (eds.). *The verb phrase in English: Investigating recent language change with corpora*. Cambridge: Cambridge University Press. 2013. xxv + 445 pp. ISBN 978-1-107-01635-4. Reviewed by **Peter Collins**, University of New South Wales.

Most of the contributions to this volume originate as papers from a symposium held at the University of London in 2009 in celebration of the fiftieth anniversary of the founding of the Survey of English Usage. These were subsequently supplemented by several invited contributions. The focus is on current, relatively short-term, change in English, an emergent research area in which the most influential publication to date is Leech, Hundt, Mair and Smith (2009). While Leech *et al.*'s work is based on the 'Brown family' of corpora, contributors to the present book avail themselves of a wider range of corpora, including a spoken corpus prepared at the SEU of which Leech *et al.* also make limited use (the Diachronic Corpus of Present-Day Spoken English, or DCPSE), the *Time Magazine* Corpus of American English, the Corpus of Contemporary American English (COCA), and the Corpus of Historical American English (COHA).

Apart from one chapter on Canadian English (by Tagliamonte), all the chapters are concerned with British or American English, or both. Most of the chapters address essential categories of the verb phrase such as the progressive aspect, the passive voice and modal verbs, but some appear more tenuously related, notably Krug and Schützler's on the nominal expression *the idea is*, Davies' and Callies' on non-finite complement clauses, and Kaltenböck's on comment clauses. One might have expected some discussion by the editors of their interpretation of the category of verb phrase, and of how this interpretation is understood to encompass the constructions explored in these chapters.

A strength of the book is its concern with questions of methodology. In fact Chapters 2–6, following the authors' Introduction (Chapter 1) and preceding the more descriptively-oriented Chapters (7–15), are claimed by the editors to be distinguishable in their illustration of a range of methodologies. The most explicit and extended treatment of methodological issues is offered in Chapter 2

by Aarts, Close and Wallis, and in Chapter 4 by Smith and Leech (see below). Aarts *et al.* demonstrate the importance of an onomasiological approach focusing on linguistic variation involving choices between two or more alternants. Beginning with the progressive aspect they demonstrate that (slightly) differing results are obtained if change is measured in normalized frequency counts, or as a percentage of the total number of VPs (to eliminate the possibility that changes could be simply a by-product of VP density), or within a more restricted set of variants ('drilling down' to a true set of progressive versus 'progressivisable' VPs by removing 'knock-out' contexts such as imperatives and the BE *going to* future, in order to eliminate the possibility that change is due to other sources of variation). Aarts *et al.* concede that the third, proportional, approach is often challenging and subjective. Indeed, in the case of the progressive, the difficulty of determining whether a VP is progressivisable is compounded by the recent emergence of new uses as discussed in other Chapters (4, 8, 9, 12): for example the BE + adjective construction as in *I'm being primitive* (from COCA 1996), the 'recentness' use as in *I was hearing the other day that ...* (BNC), and the epistemic comment clause use as in *I'm guessing*.

In Chapter 3 Davies demonstrates how robust data from the 100-million-word *Time* corpus (1920s–2000s) allows for more fine-grained analyses of developments with certain types of non-finite verbal complement (such as V NP *into* [V-ing] as in *We talked Bill into staying*) than would be possible in smaller corpora. For example the construction just mentioned yields a mere 29 tokens in the BROWN family quartet, and while their distribution between the two 1960s corpora (9 tokens) to the two 1990s corpora (20 tokens) is suggestive of a recent increase in frequency, this is confirmed by the distribution of the 1,101 tokens in the *Time* data, where there is an almost fourfold increase from the 1920s to the 1970s. While Davies fails to mention the newer additions to the BROWN family (BLOB-1901, BLOB-1931, BE06 and B-Brown, as discussed in Smith and Leech's chapter), the problem remains that such corpora are too small for low frequency items, and that they sample the language only every three decades. Davies' chapter is complemented by Smith and Leech's demonstration in Chapter 4 of the value of the million-word Brown family of corpora in studies of higher-frequency items. Not only are these corpora multigeneric (contrasting with the *Time* corpus, which is monogeneric and in fact based on issues of a single magazine), but they also enable comparisons to be made across the two major regional varieties of English. While recognizing the desirability of the proportional approach to measuring change, Smith and Leech defend their use of normalized (per million word) frequencies as having special relevance to "the word-counting principle (500 texts of 2,000 words each) which underlies the

comparability of corpora in the Brown family” (p. 75). A strict proportional approach – whose complexity becomes problematic as the number of alternatives grows – is also eschewed in Smith and Leech’s analysis of the multiple choices involved with the various core and emergent modals in the semantic field of obligation and necessity. They opt instead for a more loosely defined ‘ecological’ approach, and appeal to processes such as grammaticalization, colloquialization and democratization to explain such diachronic trends as the decline of *must* and the rising popularity of HAVE *to* and NEED *to*.

The placement by the editors of Biber and Gray’s chapter, ‘Nominalizing the verb phrase in academic science writing’ (Chapter 5), in the set of ‘methodological’ papers is somewhat puzzling, given its descriptive orientation and reliance on straightforward normalized frequencies. The chapter begins by presenting evidence of an increase in ‘nominal style’ and decrease in ‘verbal style’ that is a mostly 20<sup>th</sup> century development and particularly associated with academic writing. Ultimately, however, the authors raise more questions than they answer: their focus is on the numbers and types of verbs used in different historical periods in science writing, while complementary investigation of nouns (and more specifically, nominalizations) is not pursued on the grounds that this would be “outside the scope of the present volume” (p. 114).

The only chapter to systematically use apparent time rather than real time measurements of change is that by Tagliamonte, who reports findings from a project studying the speech of 275 people from Toronto, Canada, aged between nine and 92 (and offers salutary tales of caution for those unaware of the challenges in attempting to collect a random sample of sociolinguistic data). Tagliamonte’s approach, rooted in Labovian sociolinguistic traditions, uses similar proportional methods to Aarts *et al.* (Chapter 2) and other authors. These enable her to identify several shifts – the rise of statal possessive HAVE at the expense of HAVE *got*, and of deontic HAVE *to* at the expense of other modal variants – that in their extent (near categorical in the first case) suggest a distinctive Canadian trajectory. Factors in the broader sociohistorical context, such as geographical mobility, shifting social norms and technological developments, are invoked as sources of explanation for the changes documented.

Most of the remaining more descriptively-oriented chapters provide case studies of topics touched on in the previous chapters, notably progressives (Levin; Pfaff, Begs, and Hoffmann), modal expressions (Williams; Johansson), and non-finite complementation (Krug and Schützler; Callies). Additional topics are explored in the papers by Pérez (operator contraction), comment clauses (Kaltenböck), and Bowie, Wallis and Aarts (the perfect aspect).

The two chapters on the progressive deal with particular uses, Levin's (Chapter 8) with the BE *being* ADJ type as in *she is being silly*, and with private verbs such as THINK, HOPE and WISH), and Pfaff *et al.* (Chapter 9) with the use of past progressives in recent past contexts as in *I was just reading...* Levin focuses on American English using data from the *Time* corpus, supplemented by COCA and the Longman Spoken American Corpus (LSAC). While the rise of BE *being* ADJ is found to have slowed since the 1960s in *Time*, after a fourfold increase from the 1920s to the 1960s, private verbs have experienced a significant growth spurt since the 1960s. In explaining the latter development Levin echoes Smith and Leech (Chapter 4) in citing as relevant factors the systemic processes of subjectification and generalization, and the sociostylistic processes of colloquialization and democratization. The 'recentness' use of the past progressive identified by Pfaff *et al.* is, unlike the subjective uses discussed by Levin, an aspectual use (albeit resultative/perfective rather than imperfective) in which the progressive appears to be competing with non-progressive forms (compare *I was just reading* with *I just read/I have just read*). Frustratingly, while Pfaff *et al.* discuss and exemplify such alternation, noting the American preference for the simple past, and the British preference for the present perfect, their corpus searches in BNC and COCA focus solely on the progressive. The results of a grammaticality judgment survey are discussed but none of the 40 stimulus utterances – despite a claim to the contrary (p. 225) – contains a simple past verb form.

Williams' chapter on the VP in legislative English finds evidence in a set of small corpora of sweeping changes, particularly with modal verbs, undoubtedly resulting mainly from reforms driven by the Plain Language movement. The most striking of these is the demise of *shall*, traditionally a distinctive marker of legal English, which declines from 276 to zero tokens between 1980 and 2010 in Williams' British data, and from 162 to 0 in his Australian data. An increase in the frequency of *must*, *may* and BE *to*, one that runs counter to their decline in other genres, is suggested to be a consequence of the fate of *shall*. Johansson's chapter, which traces developments with four modals/semi-modals of obligation/necessity, *must*, HAVE *to*, HAVE *got to* and NEED *to*, complements Smith and Leech's discussion of these modal expressions. The recency of his COCA data (1990–2008) enables Johansson to identify, *inter alia*, a reversal in the rise of HAVE *to*, the most frequent by far of the four items. Reading this chapter will be a moving experience for anyone who knew Stig Johansson. Following his death in 2010, the editors decided to include a minimally edited version of the draft he used for his presentation, anticipating that readers "will catch something of the authentic cadence of Stig's voice, as it was heard at the symposium" (p. 372).

Krug and Schützler use data from a range of corpora to demonstrate that changes occurring with the *The idea is ...* construction (including a frequency rise, an increase in *to*-complementation, and a trend towards the marking of intentionality meaning) reflect ongoing grammaticalization. Callies uses the diachronic implications of comparisons between his BNC- and COCA-based data and earlier studies to demonstrate that *help* is renouncing its status as the only English verb that can be used both transitively and intransitively with a bare infinitival complement. His findings also support Rohdenburg and Schlüter's (2009: 6) hypothesis that contemporary American English is more prone than British English to omit semantically redundant function words.

The remaining three chapters may be considered together. Pérez finds that in spoken British English BE increasingly favours operator contraction over negative contraction (*'s not* versus *isn't*, etc.), while the other contractable operators continue to favour negative contraction (*wouldn't* versus *'d not*, etc.). Kalt-enböck explores a further stage in the grammaticalization of comment clauses such as *I think* and *I reckon* from marking epistemic modality to discourse/pragmatic markers. Bowie *et al.*'s study of the perfect in spoken British English finds *inter alia* that the present perfect, contrary to studies showing it to be in decline, is rising if tokens are measured as a proportion of tensed past-marked VPs, while other perfects are declining. Bowie *et al.* concede that their result may simply be a by-product of "the decline of TPM [tensed, past-marked] VPs in the corpus, and a concomitant increase in present-oriented contexts" (p. 335), an indication that their proportional approach may be no less immune to problems of interpretation than normalized frequency methods.

This is a cutting edge volume, excellently presented with very few errata (e.g. Davies does not in fact include the V (*to*) V construction (p. 3); the second HAVE *to* on p. 83 should be NEED *to*), which should be essential reading for scholars working on recent language change.

## References

- Leech, Geoffrey, Marianne Hundt, Christian Mair and Nicholas Smith. 2009. *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.
- Rohdenburg, Gunter and Julia Schlüter (eds.). 2009. *One language, two grammars? Differences between British and American English*. Cambridge: Cambridge University Press.