

- Conde-Silvestre, Juan Camilo. 2021. Review of Merja Stenroos & Kjetil V. Thengs (eds.), *Records of real people: Linguistic variation in Middle English local documents*. <https://linguistlist.org/issues/32.3987> (accessed 12 July 2022).
- Gordon, Moragh. 2017. The urban vernacular of late medieval and Renaissance Bristol. PhD dissertation, Utrecht University.
- Hernández-Campoy, Juan Manuel & Juan Camilo Conde-Silvestre. 2015. Assessing variability and change in early English letters. In Anita Auer, Daniel Schreier & Richard J. Watts (eds.), *Letter writing and language change*, 15–34. Cambridge: Cambridge University Press.
- McIntosh, Angus, Michael Samuels & Michael Benskin (eds.). 1986. *A Linguistic Atlas of Late Mediaeval English*. Aberdeen: Aberdeen University Press.
- Thaisen, Jacob. 2019. The round allograph of <r> in late Middle English. *Studia Anglica Posnaniensia* 53, 129–44.
- Wray, Alison. 2008. *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.

(Received 14 July 2022)

doi:[10.1017/S1360674322000272](https://doi.org/10.1017/S1360674322000272)

Erik Smitterberg, *Syntactic change in Late Modern English: Studies on colloquialization and densification* (Studies in English Language). Cambridge: Cambridge University Press, 2021. Pp. xii + 302. ISBN 9781108564984.

Reviewed by Christian Mair, Universität Freiburg

In his new book Erik Smitterberg investigates diachronic changes in four syntactic variables in the period from 1700 to 1900: *not*-contraction; co-ordination by *and*; nouns serving as premodifiers of other nouns; and participles postmodifying nouns. The first two variables were chosen as syntactic symptoms of colloquialisation, the last two as indexes of information compression (or, to use the author's preferred term: densification). The primary sources of data are the *Corpus of Nineteenth Century English* (CONCE) and the *Corpus of Nineteenth-Century Newspaper English* (CNNE). Comprising *c.* 1.3 million words in total, both are carefully compiled small corpora (by present standards of size). The corpus analyses presented in the book are traditional in the best sense of the word. They combine carefully compiled descriptive statistical surveys of corpus frequencies (with some additional multifactorial analysis) and philologically competent qualitative analysis of selected individual examples in their textual and historical context.

There is by now a large body of corpus-based research on colloquialisation and densification, and also on the specific variables in focus here. In such a situation, it is not the study design itself – solid and well thought out as it may be – that will provide the source of innovation. But there is ambition and an innovative thrust at a higher level, because the presentation of the corpus findings is complemented with much theoretical discussion that addresses some of the fundamental theoretical issues in

historical linguistics and contains deep and original methodological reflections on the role of corpus data in modelling grammatical change. One issue that is particularly prominent is the question why, in an age of ever-increasing idiolectal variability and rapid expansion of precisely those types of loose social networks that are widely assumed to promote change, corpus data seem to suggest a slowing down of the diachronic dynamic in the recent past, in comparison to what went on in Middle English and Early Modern English. This is what the author refers to as the Late Modern English ‘stability paradox’. The study also raises all the right questions about the complex relationship between text-linguistic and sociolinguistic approaches to language history, which both approach structural changes in the linguistic system through the discourse level (i.e. the changes in historically evolving traditions of speaking and writing), but do so in different ways.

Chapter 1 (‘Introduction’) briefly lays out the plan for the book. Chapter 2 (‘Sociocultural and linguistic change in Late Modern English’) explores the impact of urbanisation, industrialisation and compulsory education on the structure of social networks and contains a lot of intelligent discussion of such networks as mediators between sociocultural and linguistic changes. At the end of this chapter, the author identifies three ways of resolving the stability paradox: (i) The correlation between loose networks and faster change is not as generalisable as is widely assumed. (ii) The expansion of loose networks may not have taken place at the rate assumed. (iii) There was more syntactic change during the Late Modern English period than is recognised in the literature. As the author points out, the three lines of argumentation are not mutually exclusive. Anticipating his conclusions, however, he identifies (iii) as the scenario that accounts for most of the observed data.

Drawing on a wide range of classic and recent models of linguistic change, the author uses chapter 3 to explain the principles of his proposed resolution of the stability paradox. He distinguishes between three basic types of language change, independent innovation, propagation and propagation-dependent innovation (by which he understands imperfect replication during propagation). He points out that network density has little impact on the first type of change, but that loose networks very much promote the other two. At the level of the individual idiolect (the true locus of change for Smitterberg), the result is increased variability because of increased exposure to other idiolects and, above all, the increasingly important standard variety. In communication that reaches beyond tightly knit local networks, speakers who are exposed to and have internalised standard norms produce output that, if subsequently collected in corpora, will give the impression of more stability across time than may have been evident in spontaneous communication. Regional, social and stylistic fluctuation and volatility in myriad idiolects may thus be levelled out in corpora, which, after all, document the average usage of large communities.

Chapter 4 explains the design of the corpus study in meticulous terms. The central exploratory concepts of colloquialisation and densification are discussed in detail. The same goes for corpus methodology. In a research community in which the trend is increasingly being set by advocates of big data and ever more sophisticated statistical

procedures, the author concedes that his philologically aware traditional approach may put him on the defensive. This does not make him overly apologetic, however, as is shown by his remarks on the pointlessness of much testing for statistical significance (discussed in a euphemistically titled subsection on ‘The use of significance testing’ on pp. 105–6) and his spirited plea for the continuing relevance of ‘small and tidy’ corpora (pp. 106–7). I agree with Smmitterberg’s claim that the relationship between statistical significance and linguistic significance is complex, and I would recommend reading this chapter to anyone who needs an idea of how much linguistic-philological expertise and expert judgment are necessary to extrapolate from a frequency trend in a diachronic corpus to the corresponding language change (at the community and even more so at the individual level).

Chapters 5 to 8 present the four case studies. The results for *not*-contraction largely confirm expectations. In CONCE it is absent from the core written genres, but makes slow headway in the speech-based and speech-purposed genres (trials, drama, fiction). The author deserves credit for the depth of his qualitative analyses – for example of variation between *Would not she ...?* and *Would she not ...?* in questions, or *was not it?* and *was it not?* in tags, which include exploration of the Pitman shorthand system used by court recorders to assess the phonetic realism of their transcriptions (p. 141, footnote 7). I also agree with him in his conclusion that few of these patterns can be generalised beyond his own corpora in view of the small numbers. In the following chapter on *and*-coordination, the author distinguishes several types. Sentence-initial *and* is a straightforward colloquialism that, like *not*-contraction, was stigmatised in prescriptive grammars. Particularly interesting are the comments on ‘super-phrasal coordination’. This term is intended to alert the reader to the complex border territory between what other taxonomies categorise as phrasal and clausal coordination by *and*. Unlike sentence-initial *and*, the set of *and*-constructions discussed under this heading was not on the prescriptivists’ radar (apologies for this anachronism), but is shown to have contributed to increasing genre differentiation in writing during the nineteenth century in subtle and subconscious ways. Given the frequency of *and*, even the small corpora analysed here mostly provide enough examples for statistical generalisations.

The first case study on densification is devoted to nouns premodifying nouns, a standard and oft-researched indicator. Not surprisingly, the general trends in CONCE and CNNE broadly confirm previous research. In several places, however, Smmitterberg is able to point out interesting details. For example, mid-nineteenth-century newspapers from CNNE that were aimed at a working-class readership did not fully take part in the general trend – an interesting case of audience design on the part of the journalists and editors. What is not often found in other research on *noun + noun* sequences is the laborious semantic classification of modifier–head relations, which is a task that cannot be outsourced to POS-tagging software. Low numbers are a potential problem again in the concluding case study on participle clauses post-modifying nouns. On the other hand, Smmitterberg argues convincingly that this structural pattern serves such a wide range of semantic and textual functions that little can be inferred

from global statistical trends anyway. To prove his point, he cites an outlier text in his corpus that was produced by a science writer who still clung to a narrative mode of presentation, and hence used many restrictive past-participle clauses, at a time when his field had already moved on to a more modern, argumentative mode of presentation (p. 255). At this point, readers with less respect for the philological tradition in corpus linguistics might become somewhat impatient and point out that problems with such outliers would not have arisen in a large corpus. Interestingly, the working-class papers differ from the rest also on this densification variable.

The concluding chapter 9 summarises the findings on the major research questions raised in the study. Starting with the most abstract and general issue, the stability paradox, Smitterberg argues that there is less evidence of independent syntactic innovation in Late Modern English than in previous periods, but that this does not necessarily mean that there was less innovation at the idiolectal level. He submits that at a time when there was so much propagation going on within and across loose networks, speakers' capacities to absorb new usages were getting close to saturation point, which reduced the chances of independent innovations being taken up widely enough to show up in corpus data. A standard language bias effective at individual and community levels may have played an additional role. This scenario is plausible; however, as the author himself admits, it contains an element of speculation, which can, and should, be eliminated in further research. As regards the methodology, Smitterberg defines the conditions in which corpus data can be used to support an idiolect-based analysis of language change. To sum up a complex discussion in simple terms, Smitterberg accepts that a description with full idiolectal granularity but no generalisations is pointless and then goes on to warn that statistical trends obtained from corpus data without reference to the enormous internal heterogeneity of almost any corpus are bound to remain spurious. This is a diagnosis that is no doubt plausible and on the basis of which he defends the small-corpus approach he has adopted in his book. Indeed, his studies provide much supportive evidence showing how subcorpora, specific networks or even individual speakers and writers position themselves differently with regard to each other and the average corpus norm.

On the specific language-historical problem of colloquialisation/densification Smitterberg's findings will not require a major rewrite of the relevant portion of the history of Late Modern English, but his fine-grained analyses have undoubted merit for their high degree of granularity and the very perceptive discussion of individual examples in their context. Some points are, perhaps, made more often than necessary, but the over-all impression the book makes is to remind readers of the full complexity of grammatical changes when we study them through the lens of corpora and as part of historically evolving traditions of speaking and writing. The book shows the author's erudition, not only in his own discipline, but also in British history and culture. Given this wide scope, the coverage of the relevant literature cannot be complete, but it is certainly comprehensive.

There is one strand of research that is not referenced, but might have deserved a mention. This is work by Elke Teich and her team (e.g. Degaetano-Ortlieb, Menzel &

Teich 2019; Fischer *et al.* 2020; Menzel 2022) on the *Royal Society Corpus*. Their computational and statistical approach is very different from Smitterberg's, but there is considerable overlap with regard to the time period and some phenomena in focus, especially the growing differentiation of the scholarly register into various 'dialects' for different disciplines (cf. e.g. pp. 217–18, 255 and 267 in the present book). As for language, style and editorial matters, the study meets the highest standards. I have not found a single typo, error or inconsistency that would have had an impact on correctness or comprehensibility.

All things considered, Erik Smitterberg deserves praise for a twin achievement. For those readers who are primarily interested in Late Modern English syntax as it relates to colloquialisation and densification, his book provides rich corpus illustration and cogent analyses on many points of detail. Readers coming to the book with a more general interest in the role of corpus approaches in modelling syntactic change will benefit from much intelligent discussion of how the collective community record preserved in corpora relates to the historical dynamics of social networks and individual usage.

Reviewer's address:

Englisches Seminar

Universität Freiburg

D-79085 Freiburg i. Br.

Germany

christian.mair@anglistik.uni-freiburg.de

References

- Degaetano-Ortlieb, Stefania, Katrin Menzel & Elke Teich. 2019. Typical linguistic patterns of English History texts from the eighteenth to the nineteenth century: An information-theoretic approach. In Isabel Moskowich, Beatriz Crespo, Luis Puente-Castelo & Leida Maria Monaco (eds.), *Writing history in Late Modern English: Explorations of the Coruña Corpus*, 58–81. Amsterdam: John Benjamins.
- Fischer, Stefan, Jörg Knappen, Katrin Menzel & Elke Teich. 2020. The Royal Society Corpus 6.0: Providing 300+ years of scientific writing for humanistic study. *Proceedings of the 12th Language Resources and Evaluation Conference* (Marseille, May 2020), 794–802. Paris: European Language Resources Association.
- Menzel, Katrin. 2022. Medical discourse in Late Modern English: Insights from the Royal Society Corpus. In Turo Hiltunen & Irma Taavitsainen (eds.), *Corpus pragmatic studies on the history of medical discourse*, 79–104. Amsterdam: John Benjamins.

(Received 26 July 2022)