

Book Review

Ewa Jonsson and Tove Larsson (eds.). 2020. *Voices past and present – Studies of involved, speech-related and spoken texts. In honor of Merja Kytö.* (Studies in Corpus Linguistics 97). Amsterdam & Philadelphia: John Benjamins. ISBN: 9789027207654 (hardback), xiv, 348 pp. €99.00/\$149.00.

Reviewed by **Arja Nurmi**, Faculty of Information Technology and Communication Sciences, Tampere University, P.O. Box 300, FI-33014 Tampere, Finland, E-mail: arja.nurmi@tuni.fi. <https://orcid.org/0000-0002-6446-9356>

<https://doi.org/10.1515/jhsl-2021-0011>

The Festschrift is a difficult genre, but the volume of studies to honour the work of Merja Kytö manages to be coherent in its theme. From the foreword by long-time collaborator Jonathan Culpeper, giving an overview of the development of the field(s) Kytö has been so instrumental in developing, to the select list of publications by Kytö closing the volume, the recipient of the Festschrift is evident. Through all chapters, reference is made to Kytö's work in relevant ways, discussing corpus approaches to involved, speech-related and spoken texts from the Early Modern, Late Modern and Present-day English periods. All authors make careful and extensive use of various electronic sources and tools, showing a link to Kytö's ground-breaking work in corpus-based historical linguistics. The introduction by the editors shows the tight focus of the volume, as is common in present-day Festschrift volumes. Combining the past and the present in one volume is a way of seeing more clearly what the features of spoken, involved language are, and this creates new avenues for exploring the historical stages of language. In fact, if there is one flaw in the volume as a whole, it is the structuring according to diachrony. There are several papers approaching similar topics in different periods and presenting them next to each other would perhaps have made the links more apparent.

The Early Modern English part of the volume contains seven studies. The first of these is by Jonathan Culpeper and Samuel J. Oliver and it builds directly on a concept developed by Culpeper and Kytö (2010), pragmatic noise. The data used is the *Enhanced Shakespeare Corpus*, which is given a detailed description, including the annotation practices. The search procedures and categorisation of retrieved tokens are similarly explained in a way to allow readers to duplicate the study using another dataset. The authors discuss pragmatic noise items by play, taking into account the variety of Shakespeare's work and discover that the phenomenon is most commonly found in tragedies. If the first paper on Shakespeare could be seen as reflecting actual spoken language on some level, the

second and particularly the third are more text-related than speech-related. Also using the *Enhanced Shakespeare Corpus*, Dawn Archer and Alison Findlay use keywords as a key to character analysis, in the pragma-stylistic tradition. Rather than taking fictional characters as representatives of authentic language, the authors are looking at language as a way to create them. They, too, show exemplary care in explaining the details of their statistical method. The study explores fictional speech as fictional, remembering at all times that it does not represent authentic spoken language. The third paper on Shakespearian language is by Juhani Rudanko, who focusses on soliloquies in *Othello* and discusses semantic roles, particularly the Agent. Soliloquies are seen as speech to the audience, not to other characters in the play. Rudanko takes careful note of the lineage of Shakespearean texts, as editorial interventions across the centuries might have influenced the status of some lines (p. 49). All in all, the link to features of spoken language is rather tenuous in Rudanko's chapter.

Some of the remaining chapters on Early Modern English also take on literary source materials. Terry Walker and Peter J. Grund look at speech reporting verbs in prose fiction samples in the *Corpus of English Dialogues 1560–1760*, compiled by Kytö and Culpeper. Again, the authors provide an admirably clear account of what was included in the study and what was left out. Their interesting findings show that in their data neutral reporting verbs such as *say* are the most common, but that structural and descriptive verbs such as *answer* or *cry* increased over the period studied. In another chapter investigating fiction, in this case drama and jests, Irma Taavitsainen focusses on irony and other means of humour through an analysis of interjections, pointing out that common interjections may have genre-specific functions.

Taking a different tack, Jeremy J. Smith looks at involved language in religious controversy in the 1560s. He observes two communities of practice and makes comparisons to present-day online debates, particularly in terms of abusive language, showing how specialist “coded” lexicon was used by the opposite parties both in clearly theological (*godly*) and other terms (*joy*). Matti Peikola's chapter also investigates involved language; his data come from sixteenth-century titlepages and particularly the parts that function as advertising. Peikola gives a detailed introduction of the evidence we have for early advertising, and then proceeds to analyse the use of *you* and *thou* in his data. His thoughtful qualitative analysis makes connections to present-day consumer psychology.

The Late Modern English part of the volume consists of four papers. It opens with Claudia Claridge's study of epistemic adverbs in the *Old Bailey Corpus*, actual recorded speech. Claridge discusses the nature of the data and how reliable it is in terms of recording the exact expressions uttered in the court-room. Her results show that high-ranking men take the lead in the use of epistemic adverbs, which

seems to tie in with the spread of epistemic expressions from above (cf. Nurmi 2003: 117). Taking a different approach to the *Old Bailey Corpus*, Patricia Ronan presents an exploratory chapter tracking shifts in question strategies. Her study shows the increasing role of professionals of law in court-rooms and how the type of questions asked shifts to a more specific type. The chapter raises interesting questions for further studies, including, for example, the type of answers given to questions and taking into account the type and severity of the offence.

The focus of the volume broadens geographically, as Ray Hickey takes up the pragmatic marker *sure* in Irish English. He traces the early history of the feature from the seventeenth century onwards. Hickey approaches his data carefully, bearing in mind that trends observed in literary texts do not necessarily align with trends in spoken language, particularly if a feature has been stigmatised. Liselotte Anderwald investigates the development of the American English past participle *gotten*, looking at the ARCHER corpus. Anderwald's study does not very explicitly discuss the role of spoken language, although she points to the link between the verb *GET* and informal, colloquial language, which is often associated with speech-like genres. In fact, she shows that the form did not arise in spoken language but was rather a result of prescriptivism having unintended consequences.

The Present-day English part of the volume consists of seven studies. It opens with David Denison's chapter on explanatory *so*, a very recent change in spoken British English. His data include both spoken language and scripted spoken language, such as TV shows. Another chapter looking at recent changes in spoken British English is that by Ylva Berglund Prytz. Her study compares the original BNC spoken demographic part and the new BNC2014, exploring shifts in the use of future expressions, but also the comparability of the two corpora, an approach useful to many future users of both corpora. She finds some interesting differences, for example showing that *shall* is mostly used in questions now, while in the older BNC it had a broader scope.

Karin Aijmer explores a different type of data, looking at the *English-Swedish Parallel Corpus* and how the functions of pragmatic markers *sort of* and *kind of* can be made explicit for a researcher through an analysis of a professional translator's work. Pragmatic markers can be seen as an interactive feature in texts, so even though the corpus consists of mostly literary texts, there is a link to the overall theme of the volume. Also working with pragmatic markers, Anna-Brita Stenström looks at the development and use of *innit* in two corpora of teen language and compares those with the adult speakers found in the two versions of the BNC. Her careful analysis of the different functions of pragmatic markers shows that *innit* is gaining ground.

Sarah Schwartz and Erik Smitterberg take a look at EFL teaching materials and explore the prescriptive bias in them with regard to features of spoken, informal language. Their focus is on university level materials, but it would be interesting to see also how teaching materials at lower levels of education function. The authors use the *Corpus of Contemporary American English* as a reference corpus to compare stylistic advice given in teaching materials to actual usage. Their approach to the corpus is cautious, as is necessary when using such a huge dataset with the tagging errors inevitable given the size of it. The comparison of spoken and written language allows the authors to highlight the features more common in speech, and they make considered suggestions to improve the teaching of academic writing.

Providing a match for the Early Modern English studies on written dialogues in Present-day English, Signe Oksefjell Ebeling and Hilde Hasselgård look at the intensification of adjectives in both dialogues and narration in the *English-Norwegian Parallel Corpus*. Results show that intensification is more common in dialogues, but the authors also point out the established differences between scripted dialogues and authentic ones. The volume closes with a look at orality on the searchable web by Douglas Biber and Jesse Egbert. Their main focus is on “the extent to which written texts represent the characteristics of speech” (p. 317), which is of course relevant for both synchronic and diachronic linguistics. They start from features included in Biber’s (1988) dimensions associated with interaction and real-time production. An important point in Biber and Egbert’s chapter is the reminder that the searchable web does not include the synchronous interactions regarded as typical of online communication, as those are typically confined to the more private spheres of the web.

One of the central questions of the volume and of many historical linguists relates to the search for texts representing oral, speech-based, involved and interactive qualities. Present-day studies of such genres can make comparisons with authentic conversations and help us gauge our success with past texts. Still, when we are exploring the remaining evidence of the spoken language of the past, we need to be careful in the types of assumptions we make and choose our methods of interrogation carefully. As Walker and Grund point out in their chapter, “recreating the spoken language of the past is a complex puzzle that must be approached from various angles and with various methodologies” (p. 77). This volume brings together studies of speech-based and involved texts covering both the past and the present and allows us to recognise common features as well as differences. As such, it is compelling reading both for those interested in language variation and change of the past in all its forms and those who are particularly interested in the different representations of interactive language.

References

- Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Culpeper, Jonathan & Merja Kytö. 2010. *Early Modern English dialogues: Spoken interaction as writing*. Cambridge: Cambridge University Press.
- Nurmi, Arja. 2003. The role of gender in the use of *must* in Early Modern English. In Sylviane Granger & Stephanie Petch-Tyson (eds.), *Extending the scope of corpus-based research: New applications, new challenges* (Language and computers: Studies in practical Linguistics 48), 111–120. Leiden: Brill.