

Review of Gandón-Chapela, Evelyn. 2020. *On Invisible Language in Modern English: A Corpus-based Approach to Ellipsis*. London: Bloomsbury. ISBN: 978-1-350-06451-5.  
<https://doi.org/10.5040/9781350064546>

Arja Nurmi  
Tampere University / Finland

This volume aims to bridge earlier, mostly theoretically based research of ellipsis in English with a corpus-based study of Late Modern English. The goals of these two strands of linguistics are somewhat different, as theoretical studies discuss what is possible in a language while corpus-based studies are more focused on what is typical and what patterns of variation can be observed. These differences of purpose make the dialogue of the two approaches challenging at times, but nevertheless valuable. For topics such as ellipsis, where the bulk of earlier work has a more theoretical focus, studies such as the one at hand are of particular merit. This is something Gandón-Chapela does not always seem to see the value of herself, as she is at times almost apologetic for engaging in empirical work (p. 139). The focus on the history of English is equally valuable, as this is a phenomenon still not frequently studied in the historical stages of English, with the exception of Warner (1993, 1997), Nykiel (2006, 2015) and Gergel (2009).

The introductory chapter of the volume contains an extensive discussion of the characteristics of ellipsis mentioned in previous research. The specific focus of the volume is on Post-Auxiliary Ellipsis, divided into two subcategories, Verb Phrase Ellipsis and pseudogapping. The description of earlier points of view starts with the standard reference grammars (Quirk *et al.* 1985; Biber *et al.* 1999; Huddleston and Pullum 2002), and continues with the frameworks of Systemic Functional Grammar, Transformational-Generative Grammar and psycholinguistics.



The second chapter introduces the method, that is, corpus linguistics, as well as the data used. The corpus studied, the *Penn Parsed Corpus of Modern British English* (1700–1914; Krock *et al.* 2016), is somewhat misleadingly named by its compilers and this leads the author of the volume to alternate the terms Modern English and Late Modern English as synonyms, when in fact Modern English encompasses both Early and Late Modern English and starts from 1500. The second chapter includes a careful description of the search algorithm used, and provides useful information for anyone else intending to study ellipsis in the Penn corpora. Gandón-Chapela has also persisted in creating ways to work around the inevitable tagging and parsing errors in the corpus. This chapter is concluded by a careful description of the complex analysis schema.

The third chapter presents the analysis of the results in a wealth of detail, at times overwhelming, but at the same time valuable precisely because of the meticulous description of the variables and the results. In this chapter also previous corpus-based research on ellipsis in Present-day English is integrated in greater detail, as Gandón-Chapela compares her findings to those of other empirical studies. There is great merit in the broad scope of analysis, including grammatical, semantic and discursive variables, usage variables concerning diachrony and genre, as well as processing variables in terms of lexical and syntactic distance. The analysis scheme and the careful study of each variable brings forth new information about possible structures and their frequency. While there is ample data for Verb Phrase Ellipsis, the instances of pseudogapping are rarer. This leads to some discussion of variation that is not statistically significant, but seems to be treated as such anyway (p. 127). The significance testing of results is somewhat sporadic, and it is not always clear why there is testing for some variables and not others.

There are many smaller quibbles one might raise, from counting the archaic second person singular inflected forms (*shouldest*, *shalt*) separately from the other forms of the same verbs as licensors of ellipsis (p. 256–257), or not considering the overall frequency of various modals when looking at their function as licensors (p. 177), given that *will* and *would* are considerably more common than *must*, and the corpus would have provided this point of comparison. Similarly for connectors, it would have been interesting to know how common the investigated connectors are in the data altogether. That is, how far they are specifically connected to ellipsis and how far they are high up the list just because they are frequent (p. 135). At times Gandón-Chapela's focus seems to be more on what

is possible, that is, more theoretical, than what is common in the corpus, as she comments in great detail on the individual examples representing types only rarely attested in the data, but that is obviously a justifiable position also in a corpus-based study. One of the merits of the volume is that it clearly points out gaps in corpus-based research on individual variables of ellipsis.

There is also much merit in the wealth of detail. While it is difficult to see the wood for the trees at times, the discussion of all the aspects of ellipsis and provision of details with numerous examples is obviously highly useful for those wishing to carry out further study with different data sets. It is nice to see the author also bring new linguistic features, such as clause type, to the discussion of ellipsis. Once again, it would have been interesting to relate the frequency of the clause types with ellipsis to the overall frequency of them in the corpus, but as this would have disrupted the focus of the study and considerably added to the workload, it is perfectly understandable this avenue was not pursued.

The discussion of genres and the frequency of ellipsis in them suggests, as the author points out, that the phenomenon is typical of texts related to ‘oral’, ‘spoken’ or ‘colloquial’ language, to use the terms from Culpeper and Kytö (2010: 16). Another way of thinking about this might be to apply Biber’s (1988) dimensions, particularly in terms of involved vs. informational texts. It might well be that ellipsis is at home with linguistic features associated with involved texts, and this might help target future corpus-based studies of ellipsis towards texts representing such genres. Biber and Finegan (1997) identify particularly drama but also to some extent letters as consistently representing involved features in historical texts.

The third chapter is concluded by a summing up of the findings in terms of the different variables studied, with a focus on Late Modern English in particular. This is followed by the final chapter, which gives an overall summary of findings and lays out suggestions for further research. One final highly useful feature of the volume follows in Appendix 1, which lines out the corpus tool used, *Corpus Search 2*,<sup>1</sup> as well as the query language and its functions. Appendix 2 provides similarly useful information on the labels used for part-of-speech tagging and parsing in the *Penn Treebank* corpora. For anyone

---

<sup>1</sup> <https://corpussearch.sourceforge.net/>

not familiar with the model, these provide a useful introduction and necessary support for understanding what has been retrieved through corpus searches.

While the volume has its problems in terms of, for example, significance testing, the range of linguistic and textual variables analysed provides many potential starting points for the further corpus-based study of ellipsis both in the historical stages of English and in Present-day English. Gandón-Chapela herself suggests further studies using the *Penn-Helsinki* corpora of Old, Middle and Early Modern English, which would seem like a fruitful direction, since the corpus used in this study copies its structure from the *Helsinki Corpus*<sup>2</sup> and shares the parsing and tagging model with the *Penn-Helsinki* corpora. The clearly explained search algorithms the author has developed could be put into use very easily and the results would be comparable in a very direct way. As Gandón-Chapela's results seem to suggest that involved texts are a particularly fruitful ground for ellipsis, a study using the *Parsed Corpus of Early English Correspondence*<sup>3</sup> might provide further interesting data.

While the title of the volume is somewhat misleading, as the volume focuses on a specific subtype of ellipsis and a particular period in the history of English, there is a great deal of value in the first corpus-based diachronic study of post-auxiliary ellipsis in English. This study, even with its flaws, provides a good starting point for future research and gives us much detailed information on ellipsis based on both quantitative and qualitative analysis. The results are valuable for both diachronic and synchronic future studies, and seem to provide new information on what is possible as well as what is typical in case of ellipsis.

#### REFERENCES

- Biber, Douglas. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, Douglas and Edward Finegan. 1997. Diachronic relations among speech-based and written registers in English. In Terttu Nevalainen and Leena Kahlas-Tarkka eds. *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*. Helsinki: Société Néophilologique de Helsinki, 253–275.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman Grammar of Written and Spoken English*. London: Longman.

---

<sup>2</sup> <https://varieng.helsinki.fi/CoRD/corpora/HelsinkiCorpus/>

<sup>3</sup> <https://varieng.helsinki.fi/CoRD/corpora/CEEC/pceec.html>

- Culpeper, Jonathan and Merja Kytö. 2010. *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: Cambridge University Press.
- Gergel, Remus. 2009. *Modality and Ellipsis: Diachronic and Synchronic Evidence*. Berlin: Mouton de Gruyter.
- Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Kroch, Anthony, Beatrice Santorini and Ariel Diertani. 2016. *The Penn Parsed Corpus of Modern British English (PPCMBE2)*. Department of Linguistics, University of Pennsylvania. <https://www.ling.upenn.edu/ppche/ppche-release-2016/PPCMBE2-RELEASE-1>
- Nykiel, Joanna. 2006. *Ellipsis in Shakespeare's Syntax*. Silesia, Poland: University of Silesia PhD dissertation.
- Nykiel, Joanna. 2015. Constraints on the ellipsis alternation: A view from the history of English. *Language Variation and Change* 27: 1–8.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Warner, Anthony R. 1993. *English Auxiliaries: Structure and History*. Cambridge: Cambridge University Press.
- Warner, Anthony R. 1997. Extending the paradigm: An interpretation of the historical development of auxiliary sequences in English. *English Studies* 78/2: 162–189.

*Reviewed by*

Arja Nurmi  
 Faculty of Information Technology  
 and Communication Sciences  
 P.O. Box 300  
 FI-33014 Tampere University  
 Finland  
 e-mail: [arja.nurmi@tuni.fi](mailto:arja.nurmi@tuni.fi)