

Book Review

McEnery, Anthony & Helen Baker. 2017. *Corpus Linguistics and 17th-Century Prostitution. Computational Linguistics and History* (Research in Corpus and Discourse). London, Oxford, New York etc.: Bloomsbury Academic. ISBN: 9781472506092 (hardback), xi, 257 pp. £98.00/\$128.00

Reviewed by **Merja Kytö**, Department of English, Uppsala University, Box 527, Uppsala SE-751 20, Sweden, E-mail: merja.kyto@engelska.uu.se

<https://doi.org/10.1515/jhsl-2017-0026>

This monograph by Anthony McEnery and Helen Baker brings together a number of areas of study, among them historical corpus linguistics, computational linguistics, word studies, and social history. The authors have aimed to “explore, in a spirit of full cooperation, what a linguist and historian can achieve together” (p. 1). The study thus takes us back in time to the socio-cultural circumstances of seventeenth-century England and focusses on “women who traded sex for cash or some benefit in kind in the seventeenth century” (p. 2). The broad aim of the book is to investigate how prostitution was represented in contemporary writings: how did people think and write about those involved in the sex trade and the trade itself, and what was the language and discourse about these issues like? These are intriguing questions about the life of almost any sections of society in the past, and they may be the more intriguing when targeting those involved in controversial professional activities such as prostitution.

The seventeenth century was a period of political and social upheaval in the history of England, and it was also the period when the English language continued to take shape and navigate towards what was subsequently known as ‘Standard English’. As their source of information, the authors opted for public discourse and “general written English from the period” (p. 2), setting aside specialized documents such as court records, which might have conveyed judicial rather than general attitudes to prostitution. Clearly, the goal set for the study required advanced methodology. The methodology the authors chose comprises automated searches of quantitative data in large-scale language corpora containing hundreds of millions of words drawn from machine-readable texts of various types, and traditional close readings and qualitative analyses of examples drawn from texts. This is precisely where the skills of a historian and a corpus linguist could be harnessed to raise and answer research questions in a joint creative effort.

Chapter 1, “Introduction”, presents the aim of the book and surveys the ways in which electronic corpora, big and small, have been used in previous

research. It also assesses the potential of such sources for the purposes of the present study. Large-scale corpora rather than smaller corpora are found better suited for the study. The strengths and especially the weaknesses of resources such as Google Ngrams and the Culturomics techniques are also discussed at some length. For various reasons all these resources, including the *Old Bailey Corpus* comprising trial transcripts from the period, are found less ideal in view of the present study. The chapter then introduces the resource to be used in the study, the gigantic *Early English Books Online* (EEBO) collection, which comprised 40,000 texts at the time when the book went to press (p. 19); the 1600–1699 section of the third version of the corpus used for the study included nearly over 996 million words (p. 207, note 15). EEBO was found to offer many advantages for the envisaged study, among them the availability of text files (transcriptions and facsimile versions) for further computational processing and a part-of-speech tagged version. To conclude the chapter, the CQPWeb search engine and notions such as collocation, collocate, colligation, keyword and semantic field are introduced.

One of the innovative features that the authors developed to enhance their methodology was to define the status of collocates and semantic fields in terms of their consistency in the data across the century. Four categories were distinguished, i.e. consistent, initiating, terminating and transient collocates. To qualify as a consistent collocate, the item had to occur in at least seven decades across the century investigated (pp. 26, 28); similar criteria were defined for the other types of collocates. Methodological developments such as this add to the operationalizability and replicability of the research carried out in the study. Overall, the lucid account of corpus linguistic techniques presented in the chapter is very useful and can be recommended to all those interested in the potential of machine-readable texts for various research purposes. The authors take into account the varying interests of their audience and helpfully provide clues as to what sections to skip, should the discussion prove too technical or off the reader's interest range.

Chapters 2 and 3 are devoted to describing aspects of Early Modern prostitution as well as official and popular attitudes towards it. It is pointed out that rather than constituting a single homogeneous group, sex workers in the seventeenth century had their individual backgrounds and life trajectories. As poverty has been identified as the factor that led people to earn their living in the sex trade, Early Modern prostitutes were concentrated in certain echelons of society and were mostly illiterate women from the lower-class ranks (pp. 33–35). Understandably enough, historians have found it difficult to reconstruct the lives of these women based on their own voices. Yet there is some evidence bearing witness to the circumstances in specimens of oral

culture in the form of e.g. proverbs, folklore, ballads, charms, anecdotes and tales from the period (p. 35). To make sense of the complex picture of early modern prostitution, the authors wisely point out that they intend to draw on gender history so as to consider female prostitutes and their identities in interactions with men in their various roles as e.g. clients, facilitators or exploiters rather than in isolation (p. 37).

Chapter 2 further examines the regulation and punishment of prostitutes in Early Modern England, describing the dire conditions of life in the houses of correction at the time (pp. 46–47), and the official, mostly harsh, attitudes towards prostitutes in the capital that were manifested in legislative and executive measures taken by the kings and the Long and Rump Parliaments dominated by the Puritans. Under Cromwell's leadership, prostitutes faced particularly stern circumstances (p. 51). These sections in the book, followed by a description of the conditions that led women to prostitution, whether tricked or forced into it (in the minority of cases) or willingly entering the trade (in the majority of cases) (p. 65), make vivid reading. That women in the Early Modern period were considered weaker and less intelligent than men, and overwhelmed by their sexual passions, may not be new information to many readers, but the detailed description of the hierarchy of the various types of prostitutes and the localities where they exercised their trade is likely to make many readers gain a better picture of the complexity of societal structures and women's conditions in the sex trade. As it is aptly put in the book, "[e]arly modern women who engaged in prostitution were categorized and largely condemned by their society" (p. 89).

Chapter 3 turns to popular attitudes towards prostitutes, depicting how sex-trade workers were presented in seventeenth-century literature. The subject clearly fascinated writers in the period, who often relished in portraying whores as "skilled manipulators and liars" (p. 94). Indeed, there is evidence of how men thought about prostitutes and women in general in various sources such as broadsides, pamphlets, sermons, petitions, diaries, newspaper articles and fiction literature (p. 91). While history writers have debated how accurately these documents reflect lifestyles of the period, undeniable historical events can throw light on contemporaries' views on brothels and sex workers. A frequent response was violence, and brothels were attacked by infuriated inhabitants, especially by apprentices in London, who were dissatisfied with their employment conditions and declared their moral values by targeting prostitutes they considered dissolute (p. 99). Overall, prostitutes were often looked down on as they were considered to threaten the social order and orderly behaviour (p. 103). As explained in the chapter, a manifestation of this understanding was the establishment of the Society for the Reformation of Manners (SRM), an offshoot of

religious societies in England towards the end of the seventeenth century. This section in the book aptly throws light on the variety of fluctuating motives underlying the activities of SRM that eventually led to its decline and failure to achieve its original goals (pp. 107–109). The ensuing sections on links between prostitutes and crime offer case studies from real life and literary sources, painting a lively picture of the repertoire of alleged criminal activities among sex workers and closing with an assessment of whether prostitutes actually formed part of a ‘criminal underworld’ as has been claimed in previous literature.

While Chapters 2 and 3 provide a background to the study of public discourse and how prostitutes were represented in it, Chapters 4 to 6 account for the results of the research carried out on EEBO and the other sources consulted. The way to operationalize the aim of the book into hands-on work was to investigate words used to refer to prostitutes in the period. This may sound like a simple task on the face of it, but, as the authors show, it involved having to deal with a number of difficult methodological choices so as to justify the data selection and further processing. As appears in Chapter 4, it is not off-hand clear what words should be included in the searches intended to cover the terms used to indicate a prostitute in the period. The authors investigated various scenarios to perform this task, among them the use of their own intuitions about the past language form, the terms used by historians in their writing, and consulting present-day and contemporaneous dictionaries and lexicons. The discussion of these options highlights a number of the difficulties in linguistic work which, even though it can be facilitated enormously by computerized search techniques and language corpora, nevertheless requires labour-intensive human effort.

Among the time-saving techniques is the use of collocations and part-of-speech tagging as guides. For instance, it appears in the discussion of the words used by historians that the word *traffic* as a noun with 10,573 instances in EEBO in 3,221 texts (p. 223, note 20) collocates with some 20 terms (e.g. *Commutative*, *Commerce*, *merchandising*, *Navigation*, *sea-faring*) with reference to ‘a trade in goods for profit’ and that one of the collocates, *infamous*, as in *infamous traffic*, refers to any type of immoral transaction; four of these instances in EEBO clearly refer to prostitution (pp. 132–133). This example illustrated the benefits of the study of collocations as a gateway to shifting word meanings in the data. The method also works on high-frequency terms used by historians and listed in dictionaries, among them *lady*, *girl*, *woman* and *person* (including their plural forms). The list of expressions involving *woman* or *person* in the sense ‘engaged in extramarital or paid sex’ includes the singular and plural forms of items such as *debauched woman*, *lady of pleasure*, *lewd woman*, *lewd person*, *wanton lady*, *woman of the town*, to cite the better-represented data (pp. 137–138).

To test the value of early dictionaries as a source of data for their study, the authors explore *Lexicons of Early Modern English* (LEME), a collection of machine-readable early dictionaries and lexicons. Among these sources are dictionaries listing words and definitions for those words, and dictionaries listing ‘hard’ words, i.e. words that were new to contemporary language users or that were associated with special registers, e.g. slang. The hard-word dictionaries proved more useful than the other types of dictionaries as they gave near-synonyms in the definitions of a word. In this way, although *strumpet* is not listed as a headword in the dictionaries, the term *courtesan* in Elisha Coles’ *English Dictionary* (1676) yields the definition “a Court Lady, also a Strumpet” (pp. 141–142). The discussion in this section is very useful, as it helps the reader to follow the reasoning one needs to pursue when exploring the semantics of words in earlier periods. An example of a successful line of reasoning concerns the status of the word *whore*: this word does not usually appear as a headword in seventeenth-century dictionaries, and it is frequent in EEBO, which led the authors to argue that the word is “a strong candidate for the prototypically salient word used to refer to the concept of prostitution in this century as other near synonyms are frequently defined in relation to it” (p. 143). Yet in many instances, there was no support in EEBO for a dictionary definition, as was the case with, for instance, the word *CAT*. This very frequent word in EEBO only occurs with collocates relating to felines (e.g. *dogs*, *rats* and *mouse*). Hence, there is no evidence of its association with prostitutes in public discourse, as suggested in a dictionary entry from the 1690s (for this decade the top five collocates included *rats*, *mice*, *civets*, *mouse* and *monkey*) (pp. 144–145). On the whole, early dictionaries did not prove an ideal resource to base the present investigation on: the definitions described do not always tally with corpus evidence and may display the dictionary compilers’ inaccurate or idiosyncratic intuitions (p. 146).

Instead, the frequency data drawn from EEBO per each decade in the seventeenth century pointed to a number of promising lines of investigation. The words included in the analysis were COURTESAN, HARLOT, JADE, JILT, NIGHTWALKER, PROSTITUTE, PUNK, QUEAN, STRUMPET, TRULL and WHORE (the small caps representing lexemes covering spelling variants, and singular and plural uses). A number of interesting results emerged pointing to the need for further research. For instance, the use of WHORE peaked in the 1640s (as a proportion of all the occurrences of the eleven words included in the analysis); the following decade it was HARLOT that peaked in the data (p. 151). While JILT, PROSTITUTE, PUNK and WHORE gained ground towards the end of the century, COURTESAN, HARLOT, JADE, NIGHTWALKER, QUEAN, STRUMPET and TRULL declined (p. 151). The use of correlation statistics pointed to interesting relationships between the words in terms of

the use made of them in the period. The three strongest correlations were JILT/PROSTITUTE (JILT rose as PROSTITUTE rose), STRUMPET/WHORE (as WHORE rose, STRUMPET declined), and HARLOT/PROSTITUTE (as HARLOT declined, PROSTITUTE rose) (p. 154). Chapters 5 and 6 in the book focus on these words so as to investigate shifts in their usage and the ways in which these shifts might reflect socio-cultural changes across the period.

Chapter 5 is devoted to the words STRUMPET and WHORE, and Chapter 6 to a further discussion of WHORE and three related words in the data, HARLOT, PROSTITUTE and JILT. In addition to discussing changes in the frequencies of these items, the collocates of the words were placed in semantic fields on the basis of manual semantic analysis. The plan to use an automatic semantic tagger based on the USAS system could not be pursued owing to the discrepancies in the worldviews reflected by the system and the data. This was a wise decision, as it is true that it would have been difficult to base any definitive conclusions on the results produced by a classification reflecting semantic fields characteristic of the nineteenth and twentieth centuries when the data for the study was drawn from the seventeenth century. The definitions for the semantic fields distinguished in the analyses were developed by the authors as original research when screening the collocates in combination with the words studied (pp. 163–165).

Of the words studied in detail, WHORE emerged as a particularly complex case, not least owing to the fluctuations in its keyword status across the century and the collocation “Whore of Babylon” used as an insult during the Civil War (pp. 162–163). Accordingly, this term receives a substantial discussion in Chapters 5 and 6. The semantic fields associated with WHORE included e.g. Criminality, Gender Identity, Insult, Purchase, Family Relation, Religion, Acting, Controllers, Honesty, and Undesirable Characteristic (all these consistent across the period; for the full repertoire including initiation and transient fields, see Table 5.2, p. 167). The authors found that apart from uses related to religion, the word stood for “a consistent representation of real, rather than allegorical, whores” and even when a positive attribute was in question (e.g. Honesty), the use appeared in negated contexts (p. 168). The detailed discussion of the semantic fields of WHORE makes fascinating reading, as parallels presented by historians are brought in to accompany the results obtained in corpus-linguistic analyses.

While the collocation analyses in Chapter 5 targeted consistent collocates persisting within a field across time, Chapter 6 addresses the initiating and transient collocates for WHORE, and throws further light on HARLOT, JILT and PROSTITUTE. Regarding WHORE, the authors present convincing evidence of a change in the semantics of the word: a major shift took place in the 1660s,

the semantics of the term expanding to cover new fields, e.g. Age, Pity and Title; of these, Pity was attested only in the 1660s–1680s. The authors account for this and other corresponding results with reference to findings presented in terms of socio-cultural developments, including changes that could be postulated to have taken place in people's attitudes to prostitutes over the decades in question (pp. 186–187). In the analyses on HARLOT, JILT and PROSTITUTE, comparisons were made with WHORE and the other items. Of the words studied, JILT posed the greatest number of problems of the terms investigated. It appeared in records in the 1660s and became better represented only in the 1690s. For the less-represented decades, the authors turned to qualitative analyses. The conclusion drawn by the authors was that while JILT emerged as a lexical innovation, it was a close synonym of WHORE and other related words (pp. 194–195). The chapter concludes with an intriguing section devoted to the geographic distribution of the mentions made of PROSTITUTE, WHORE, HARLOT and STRUMPET in the data. As collocation analysis was of no help here, it was necessary to expand the methodological scope of the study and resort to a method called concordance geo-parsing. The results of this analysis, reported in another study by the authors, showed that there was a link between London and the discussion of prostitution in the corpus (p. 197). It would have been of interest to relate these findings to historical socio-cultural evidence, or if there was no room for proper discussion in the chapter, at least to offer cross-references to suitable places in Chapters 2 and 3.

Chapter 7 revisits the relationship between public discourse and historical events and surveys the methodology applied in the study. It also assesses the forms of collaboration undertaken by the linguist and the historian and encourages future interdisciplinary projects using large-scale electronic corpora.

In sum, the book stands out as a creative and carefully crafted project which highlights the potential of corpus linguistic methodology and the benefits of investing in exploiting language corpora from the interdisciplinary perspective. The need for collaboration between linguists, historians and experts in other fields has already been emphasized in many recent corpus compilation projects (e.g. the *Corpus of Early English Medical Writing* project, see <https://www.helsinki.fi/en/researchgroups/varieng/scientific-thought-styles-the-evolution-of-english-medical-writing>), and on all accounts *Corpus Linguistics and 17th-Century Prostitution* bears witness to the relevance of that call.