Corpus Pragmatic Studies on the History of Medical Discourse

EDITED BY

Turo Hiltunen

Irma Taavitsainen

p&b

John Benjamins Publishing Company

Corpus Pragmatic Studies on the History of Medical Discourse

Pragmatics & Beyond New Series (P&BNS)

Pragmatics & Beyond New Series is a continuation of Pragmatics & Beyond and its Companion Series. The New Series offers a selection of high quality work covering the full richness of Pragmatics as an interdisciplinary field, within language sciences.

For an overview of all books published in this series, please see
benjamins.com/catalog/pbns

Volume 330

Corpus Pragmatic Studies on the History of Medical Discourse
Edited by Turo Hiltunen and Irma Taavitsainen

Corpus Pragmatic Studies on the History of Medical Discourse

Edited by

Turo Hiltunen
Irma Taavitsainen
University of Helsinki

John Benjamins Publishing Company

Amsterdam / Philadelphia

∞™ The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI z39.48-1984.

Table of contents

Part II. Changing functions, roles and representations

Preface

This corpus pragmatic volume is a result of an international collaboration which originated in *the Second International Conference on Historical Medical Discourse* (CHIMED-2) in Helsinki in June, 2019. The collection brings together original contributions based on papers read at the conference, all of which share a common theme: they rely on linguistic corpora to obtain new insights into the pragmatics of historical medical discourse. By drawing on a variety of ideas and methods in pragmatics, corpus linguistics, and Digital Humanities, they illustrate the diversity and potential of corpus pragmatic approaches in understanding medical texts from the past, as well as their authors and their readers.

We would like to thank all authors for their work towards this shared goal. We are indebted to Dr. Carla Suhr and Ms. Noora Kumpulainen for their help in the early stages of the project. We are also extremely grateful to the colleagues who agreed to act as anonymous reviewers of the individual contributions, as well as the two external reviewers who commented on the entire manuscript in its finished form. Their insightful comments greatly improved the volume. We would likewise want to thank the series editor of *Pragmatics & Beyond New Series*, Anita Fetzer, for accepting the volume to the series, and Isja Conen and the editorial staff at John Benjamins Publishing Company for their valuable work.

Helsinki, 5 March 2022,

Turo Hiltunen   Irma Taavitsainen

CHAPTER 1

Corpora, pragmatics, and historical medical discourse

Turo Hiltunen and Irma Taavitsainen
University of Helsinki

1. Introduction

The register of medical of writing[1] is well-suited for historical and diachronic stud-
ies, as it has a wide scope with both professional and lay texts and has undergone
major alterations resulting from changes in the context of production, reception
and use of texts. In recent years, corpus-linguistic analysis of medical writing has at-
tracted increasing attention and the area of research is active and growing, as shown
by a number of recent publications on textual interaction (Gotti and Salager-Meyer
2006, Gotti et al. 2015) and linguistic variation between and within registers and
genres of medicine (e.g. Biber and Finegan 1994, Nwogu 1997, Salager-Meyer et al.
2007, Millar et al. 2013, Dorgeloh 2016);[2] the evolution of professional discourse
has also become a more central topic for Language for Specific Purposes (LSP; see
Swales 2000: 60).

   Research in corpus linguistics has become increasingly interested in
discourse-level questions alongside questions of lexis and grammar, and mixed
approaches drawing on both quantitative and qualitative research traditions have

---

1.   The terms *register, genre* and other terms referring to situational text varieties have been
subject to much debate in different areas of linguistics (see e.g. Ferguson 1994; Biber and Conrad
2019), and there is variation in the usage of these and related terms both in general (Conrad 2015)
and across the present volume. In this chapter, we follow Biber and Conrad (2019) in treating
the terms as different perspectives rather than different categories *per se*: thus, the term *register*
is used to refer to broad functional associations between linguistic features and communicative
purposes/situational contexts in medical writing, whereas the term *genre* refers to more specific
linkages between individual text categories and their associated features.

2.   For overviews, see e.g. Piqué-Angordans and Posteguillo (2006) and Ordóñez-López and
Edo-Marzá (2016) on present-day medical discourse, and Atkinson and Valle (2012) and Taavit-
sainen and Hiltunen (forthcoming) on historical studies; for the pragmatics of spoken English
in medical settings, see e.g. Martin (2014).

become popular (Marchi and Taylor 2018). At the same time, the field of pragmatics has become increasingly inclusive with greater emphasis placed on societal aspects of language use, following a general "pragmatic turn" in linguistics (Taavitsainen and Jucker 2015, see also Traugott 2008). This is shown, for example, in Andreas Jucker's definition of historical pragmatics, which states as its goal "to understand the patterns of human interaction (as determined by the conditions of society) of earlier periods, the historical developments of these patterns, and the general principles underlying such developments" (Jucker 2008: 894).[3] These developments have led to the emergence of an area of research known as **corpus pragmatics**, which, to quote the editor of the eponymous journal, is "the science that describes language use in real contexts through corpora" (Romero-Trillo 2017: 1). Corpus pragmatics thus views corpora as essential to understanding human communication. It is also the area that is represented by the thirteen contributions to the present volume.

The synergy between corpus linguistics and pragmatics is fruitful in the context of historical medical writing. The use of corpora enables the analysis of patterns across a wide spectrum of texts and brings quantitative rigour to the analysis of discourse (e.g. Baker 2006). In turn, pragmatics has since its inception emphasised the importance of context in all language use, and this applies equally to the study of contemporary and historical medical texts. As for the latter, cultural background knowledge and external factors are crucial to understanding past contexts and examining how meaning is negotiated in texts (Taavitsainen and Fitzmaurice 2007: 13).[4] Blommaert (2005: 17) accords an even more pivotal role to context, arguing that the most fundamental problem of discourse analysis is accounting for the relationship between linguistic forms and their *context*, encompassing all aspects of the world in which they operate. While Blommaert's focus is specifically on Critical Discourse Analysis – evident in his emphasis on social structure and the power dynamics in interactive situations such as the doctor-patient communication (2005: 39) – the general point about the importance of context is easily generalisable to all pragmatically-oriented enquiry.

In this chapter, we introduce the overall research goals of this volume and highlight some ways in which combining the two strands of research – corpus linguistics and pragmatics – can provide new insights into the history of medical

---

3. Cf. Mey's definition of pragmatics as the study of "the use of language in human communication as determined by the conditions of society" (Mey 2001: 6).

4. This idea is also in accordance with the European tradition of pragmatics that emphasises contextual analysis, including both text-external and text-internal aspects; it stands in contrast to the narrower component view of pragmatics, associated with the Anglo-American tradition; see Verschueren (2013) for a critical discussion of these terms.

discourse. We also provide a brief overview of the historical contexts within which medical texts were written across centuries, as they are crucial to understanding communicative intentions and how they are expressed and understood through the medium of language. We also survey some materials available for researchers in this field and discuss the type of annotation systems and other features facilitating pragmatically-oriented research. We will conclude the chapter with some thoughts on future directions in this field.

2.   Main themes in the present volume

The present volume represents a snapshot of current work in the corpus-pragmatic analysis of historical medical discourse. The twelve original studies that follow this introduction have a twofold purpose: they aim to improve our understanding of medical discourse in historical contexts, and at the same time test new methodologies for achieving this and critically reflect on them. The central themes running through this collection can be crystallised into three critical questions which are at the forefront in this area of research:

–   How does Western medical writing[5] change alongside societal and cultural change?
–   How can corpus evidence contribute to our understanding of medical discourse from different historical periods?
–   What are the current trends and best practices in corpus-pragmatic analysis, and what are desiderata for future work?

The chapters in this book aim at providing answers to the above questions by analysing digitally available research data and by considering relevant contextual factors when interpreting the findings obtained. Some chapters adopt a qualitative perspective characteristic of many pragmatic studies, focusing on particular texts, authors or medical topics, or specific functionally-defined discourse forms such as narratives (see **Plastina, this volume**, and **Ollikainen, this volume**). Their objective is to obtain a thorough understanding of the object of study and provide "thick" descriptions (Geertz 1973) of their linguistic aspects and their social embedding. To achieve this, they investigate specialised corpora and datasets based on a specific sub-discipline of medicine (e.g. gynaecology and obstetrics), sub-register of writing (e.g. herbal medicine, consultation letters), or themes that have repercussions on human life (e.g. ageing).

---

5.   By "Western medical writing", we refer to texts, genres, and discourses that have their origins in Classical Antiquity and Greco-Roman practices, and their subsequent, sometimes divergent developments in the Middle Ages and beyond.

Other chapters depart from broader questions of variation and change, and investigate them by using big and diverse corpora that cover long time periods up to several centuries (**McEnery and Baker; Schneider; Menzel; Hiltunen; Walker and Kytö; Schnelle et al., this volume**). Using advanced methods of corpus linguistics and Digital Humanities,[6] these studies accumulate robust quantitative evidence for linguistic and textual change. However, the role of contextualisation is equally important in these studies, as it enables researchers to interpret the trends identified through partly automated text analysis. The necessity of this stage of research is accordingly demonstrated by detailed attention to various levels of the context: the linguistic context (or cotext); the interactive context within which the texts were produced, read, and used; and the broader sociocultural and historical context.

A corpus-pragmatic perspective is a good fit for the analysis of medical discourse from a historical perspective for several reasons:

– Combining theories and ideas from pragmatics with corpus linguistic methods have already proved fruitful in many areas (see the previous section), and such a synergistic relationship has great potential to inspire new insights into historical medical discourse. The breadth and variety of data and the range of research questions calls for interdisciplinary collaboration between not only linguists and historians, but Digital Humanities practitioners of various types (cf. Rockwell 2013). While the debate between different areas of enquiry is not always easy, some steps towards that direction have already been taken in the study of historical medical discourse,[7] and the present volume also represents an attempt to bring together scholars from different parts of academia (see e.g. **McEnery and Baker, this volume**).

– Medicine is a key area of enquiry and has a long history in the vernacular continuing up to the present day. Owing to this, we have relevant data from various sources that enable the study of trajectories of discourse development across centuries – see e.g. **Schneider (this volume)**, for an insightful example of how large amounts of data can be accumulated from different sources. Still, despite the proliferation of digitised texts, the study of discursive changes in the field of medicine is not without challenges. Individual research questions frequently require researchers to collect their own specialised corpora and datasets instead of relying on existing work (e.g. **Rovelli, this volume; Grego, this volume**). The contexts of treatment, enquiry, and dissemination of knowledge change, as does

---

6. Like Mehl (2021), we recognise the close methodological affinity between corpus linguistics and Digital Humanities and share his view that corpus linguistics is, and should continue to be, one of the main pillars of DH.

7. For example, Taavitsainen and Hiltunen (eds., 2019) and Taavitsainen et al. (eds., forthcoming) include contributions by linguists, historians and data scientists.

medical terminology: new terms are introduced, some terms fall out of use, and the meanings of individual terms might change on a par with new scientific discoveries. All this presents challenges to corpus analysis and the management of data, which can be mitigated, for example, by qualitative micro-analyses, or annotation and metadata (see e.g. **Menzel**, **this volume**).

– Medicine is a specialised field, where the range of activities, functions and meanings expressed in texts and discourses are limited compared to language at large. This facilitates the analysis of individualised "speaker" meanings (cf. Levinson 1983: 17). However, an additional challenge when dealing with historical texts is the fact that the contexts of writing are less accessible to us than to contemporaneous readers, and the forms of writing from past centuries are less standardised and routinised compared to Present-day English (PDE) genres. Therefore, the importance of knowing the historical background of the texts in focus cannot be overstated.

The contributions present a range of new findings that contribute to the understanding of the interrelationship between medical texts, their writers and readers, and the context of production. Through empirical analysis of a wide range of texts, they demonstrate (1) how sociohistorical, cultural and disciplinary changes in medicine are reflected as linguistic, stylistic, and terminological changes in medical texts, (2) how data-based methods can be used to identify specific instances of linguistic/discursive change driven by cultural change, and (3) how findings obtained by quantitative analysis can be verified using qualitative analysis and triangulated against evidence from the history of science and medicine.

Given the importance assigned to context, we provide a brief overview of the historical background of Western medical writing,[8] before exploring the interfaces and synergy between pragmatic analysis and corpus data specifically in the context of English historical medical writing.

3.   Medical writing from a historical perspective

In many ways medicine is a key academic discipline with an exceptionally wide scope. It is a field that combines theoretical knowledge obtained from physical and natural sciences with practical applications and skills based on experience in the prevention, diagnosis and treatment of illnesses. The special status of medicine is also evident in the fact that its role has always been visible in people's everyday lives. The body of Western medical literature includes texts addressed both to

---

8.   For a more comprehensive account, see Taavitsainen and Hiltunen (forthcoming).

professional and lay readers on basic human concerns and needs: not only dealing with how to treat and cure illnesses but also with how to lead a healthy life, avoid falling ill and prevent the spread of disease.

From a historical perspective, medical knowledge is communicated and debated through the medium of language. External changes in the ways of doing medical research and treating patients are reflected in the language use of medical literature. Past practices and activities are mediated to us through texts, images, charts, and other multimodal elements, and they are of major importance for tracking changes in the underlying philosophy of science and medicine as well. The Western medical tradition begins in the Antiquity with Greek and Latin texts, and the history of this register in vernaculars like English and German[9] spans a period of about a thousand years from the Middle Ages to the present, with changing traditions for different audiences. Yet conceptions of science have changed, as in the earlier periods the boundaries between science, pseudoscience and magic were fuzzy or non-existent. For example, medieval treatments comprised alchemical healing potions, herbals referred to magical properties of plants, and charms mixed with religious elements. Astrology, too, belonged to the core of medicine until the Renaissance,[10] losing its status only in the seventeenth century.

The earliest vernacular texts in Britain date back to the Old English period, but they are few in number.[11] The vernacularisation of learned texts began in the fourteenth century when texts written in dialects or the incipient Central Midland Standard appear (Voigts 1984, 1989; Taavitsainen 2000; Honkapohja 2017). Some were translated or modified from canonical texts written in Classical languages, but new compositions were also created (see Pahta and Taavitsainen 2004: 13–14 and Pahta 1998). Latin prevails as the language of medical writing until c. 1700 (Webster 1975).

Medical writing evolved as a response to changes in the styles of thinking, but changes took place gradually rather than abruptly, and the medieval trends continued for centuries beyond their heyday. Late medieval and early sixteenth-century texts were characterised by scholastic argumentation that prevailed longest in texts for large and heterogeneous readerships, such as broadsheets and printed pamphlets (Taavitsainen 2018; see also **Schneider, this volume**, and **Whitt, this volume**). In learned texts, however, medieval scholasticism was gradually replaced by more systematic empirical observation, following the ideas of influential natural

---

9.   See also **Schnelle, this volume**.

10.   On the importance of astrology in 16th-century Paracelsian medicine, see Wear (2000).

11.   Prognosticatory texts and recipes are extant as well as a veterinary tract and computational writings with calculations of time.

philosophers and scientists, like Francis Bacon (1561–1626), Robert Boyle (1627–1691) and Isaac Newton (1642–1727), and medical doctors, like William Harvey (1578–1657), Thomas Sydenham (1624–1689) and William Cullen (1710–1790). The transition to empirical observation as the primary mode of knowing started at the turn of the seventeenth century (see Taavitsainen and Pahta eds. 2010; Taavitsainen and Pahta eds. 2011). From the mid-seventeenth century onwards, the cutting-edge developments took place in the activities of the Royal Society (RS), established in 1662, with participation of a varied group of gentlemen scientists. A considerable proportion of its membership were physicians who reported in writing on their explorations conducted before eyewitnesses at meetings of the society (Atkinson 1999, Hiltunen 2010). The new discourse community were eager to promote the goals and functions of the RS by developing novel ways of doing and advancing science, encapsulated in publications like the *Philosophical Transactions of the Royal Society* (PT), the first scientific periodical established in 1665. The same style is found more broadly in other writings by the members of the early RS.

The emergence of English empiricism provided a strong incentive for renewing the ways of writing. Some institutional guidance was given from above in the often-quoted instructions on the appropriate style to be adopted when writing about experiments. Another factor influencing writing style was the concurrent emergence of "polite society" (Brown 2011; Taavitsainen 2019). The establishment of the PT also meant that news about scientific discoveries were disseminated more quickly, and the new, collaborative ways of designing and performing experiments before an audience gave rise to new genres, like the experimental report (Bazerman 1988; Gotti 2001, Moessner 2008). But alongside these innovations, many traditional types of publications (e.g. treatises, handbooks, regimens, recipe books) continued to exist, and they in fact form the large majority of medical literature from this period. Another boost to scientific enquiry was brought about by the development of new instruments, such as the microscope and the air pump, which enabled natural philosophers to obtain more accurate knowledge about nature and the human body. The rise of the consumer society in the late seventeenth and eighteenth centuries extended the genre repertoire towards new directions in advertisements and medical book reviews in PT and other channels (Gray et al. 2011, Taavitsainen 2021).

Change is the prerequisite for advancement, and external changes in the ways of doing medical research and treating patients are reflected in the language of medical texts. For example, learned scholastic discourse had favoured the impersonal mode, whereas much of medical writing in the early modern period is characterised by an author-centred discourse (see **Ollikainen**, **this volume**), which was particularly characteristic of experimental reports. However, these descriptions are best seen as tendencies rather than categorical statements, as historical medical texts are rich

in intratextual variation. For example, even though the Royal Society placed great importance on neutral reporting of observations, the reports also contain rhetorical expressions of exalted feelings inspired by new observations with the microscope. Such bursts of emotion are, however, exceptional.

Eighteenth-century medical discourse can be associated with the emergence of "enquiry" as a thought-style (Taavitsainen et al. 2019), and the Late Modern English period more broadly is characterised by disciplinary specialisation, increasing professionalisation, and advances in technological developments. Statistical information is increasingly used in medical argumentation from this period onwards, beginning with the inoculation debate in the 1720s in which James Jurin, the main protagonist in favour of the new practice, made calculations in support of his views (Jurin 1724; see Taavitsainen 2020). Textually and linguistically, the eighteenth century represents a continuation and consolidation of innovations rather than radical change, and the period can be seen as paving the way to the nineteenth-century developments and even beyond (see e.g. Hiltunen 2021).

Technological advances set the scene for the establishment of modern medicine in the nineteenth century. During this period, the volume of medical titles increased dramatically (Bynum 1994), and this trend continued in the twentieth century with the emergence of evidence-based medicine and the statistical analysis of data obtained from randomised control trials. Linguistically, the mid-twentieth century represents a turning point: English started to win over French and German as the *lingua franca* of science and gradually acquired its present status. This development was unforeseen, as is the extent to which the new electronic means of communication have changed publication practices and speeded up the dissemination of findings. Today, medical articles and reports are as a rule authored by research teams (see also **Menzel**, **this volume**), and they are delivered digitally to the global discourse community.[12] The rapid dissemination of information and data sharing have of course become even more crucial due to the COVID-19 crisis, and this has boosted the popularity and visibility of preprint archives like *medRxiv*,[13] which publish reports before they have been peer-reviewed.[14]

---

12. Already in 1997, the leading medical journal *The Lancet* began to guarantee the editing and publication of "practice-changing research" within 4 weeks of submission (see https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(15)60218-5/fulltext).

13. See https://www.medrxiv.org/.

14. But see Redden (2020) on the potentially problematic nature of preprint archives.

4.   Insights from corpora

In this final section, we turn to the main issues in the corpus-pragmatic analysis of historical medical discourse in the light of the studies in this volume. The field of *corpus pragmatics* is characterised by the integration of key techniques of corpus linguistics and pragmatics, respectively. Historically, these fields of enquiry have often been presented as having contradictory aims and even seen as incompatible (Jucker and Taavitsainen 2014): corpus linguistics has been regarded as a method for discovering general patterns of language use while pragmatic methods have been seen as tools for investigating individual speaker meanings and local textual functions.[15] However, the recent years have seen an increasing degree of convergence between these approaches, motivated by disciplinary developments in both. As previously discussed, pragmatics has become more inclusive and societally aware in recent decades, and corpus linguistics has likewise begun to accommodate sophisticated and nuanced theories informed by work in the neighbouring fields. Technological advances in the identification, annotation, and exploration into discourse-oriented phenomena render corpora more useful for pragmatic analysis. As a result, we are witnessing the consolidation of a "new burgeoning discipline" (Rühlemann and Aijmer 2015: 23), which has also been demonstrated to be useful in tracing pragmatic developments in a diachronic perspective (Jucker and Taavitsainen 2014: 22).[16] The present volume further demonstrates how a corpus-pragmatic framework can shed new light on the history of a specific area of writing, medical discourse.

The use of corpora is a common denominator for all studies in this volume, but their role in individual chapters varies. As outlined above, we can distinguish two main orientations, and they roughly correspond to the two main divisions in this book. The chapters in *Part I: Tracing discursive changes* mainly use corpora in a data-driven fashion to identify trends in large amounts of data without preconceived hypotheses. The corpora used range from large to very large, and they are queried with the help of advanced computational tools. Quantitative investigation is followed by qualitative analysis. By contrast, the chapters in *Part II: Changing functions, roles and representations* depart from a specific hypothesis, theory, or discourse function

---

15. This view of corpus linguistics and pragmatics as fundamentally different types of enquiry is captured in Tognini-Bonelli's (2001) illustration of the differences between "reading a text" and "reading a corpus".

16. For overviews of developments in corpus pragmatics, see e.g. Romero-Trillo (2008), Jucker, Schreier, and Hundt (2009), Clancy and O'Keeffe (2015), and Jucker (2018) on using corpora in pragmatic analysis, and Jacobs and Jucker (1995) and Taavitsainen and Jucker (2010) on the development of historical pragmatics.

derived from previous scholarship, and provide a rigorous empirical analysis of it using corpora. The corpora and datasets in *Part II* tend to be smaller than the ones used in *Part I* due to a more labour-intensive process of data selection.

Quantitative and qualitative approaches are mutually complementary, as they shed light on different aspects of historical medical discourse. Both approaches have a somewhat different set of issues and challenges; below we list six prominent ones, which are not limited to the study of medical discourse but are present in all corpus-based analyses of pragmatics and discourse:

i.   How to establish validity and reliability for pragmatic analysis?
ii. How to ensure that recall and precision are sufficiently good?
iii. How does corpus design influence findings?
iv. How to deal with errors and "bad data" (cf. Labov 1972, Nevalainen 1999), which are ubiquitous in digitised historical archives?
v.   How to combine quantitative and qualitative analysis?
vi. How to combine findings from corpus-pragmatic analyses with insights from other fields, such as medical and social history, history of ideas, and science studies?

Related to issues (i) and (ii), identifying pragmatic features in a corpus is typically more subjective than in syntactic or morphological studies, and some of their manifestations may not correspond to any surface features of the text (e.g. Kohnen 2009). Due to this, word-based searches tend to fall short if the goal is to account for the intricate relationship between linguistic forms, communicative intentions and their interpretations in specific social contexts. One solution is to systematically add metadata and pragmatic annotation (for an overview, see Archer et al. 2008) into the corpus, a solution that is adopted in some chapters in this volume (see **Menzel**, **this volume**, and **Schnelle et al.**, **this volume**). Methodologically this is the preferred approach, as it provides a permanent record of the analysts' functional interpretations amenable to systematic computer-assisted analysis, scrutiny and subsequent expansion. Yet pragmatic annotation is also extremely resource-intensive, as it cannot be fully automated, and depends on focused individualised research agendas that may not easily transfer to other settings (Archer et al. 2008: 637). While pragmatically annotated medical corpora are still few, many available medical corpora contain metadata on text categories, which have proved useful in pragmatically-oriented work by offering information at a higher degree of granularity (see e.g. **Walker and Kytö**, **this volume**).

Corpus-pragmatic research on historical medical discourse commonly follows an approach whereby relevant concordance lines are manually annotated following some scheme deemed appropriate to the purpose at hand (see e.g. **Hiltunen**, **this volume**). Such "problem-oriented tagging" (see Smith et at. 2008) clearly offers

a useful and flexible way of identifying and describing the relevant patterns of variation. Ideally, manual annotations should be thoroughly tested and validated through an interrater validity test, yet this has not always been the practice in the past. Instead, annotations have tended to remain exclusively in the possession of their authors, which means that they do not fully contribute to the reproducibility of research (see Flanagan 2017). The adoption of standardised conventions of marking up documents using e.g. TEI-compliant XML[17] (Text Encoding Initiative; Extensible Markup Language) makes it possible to add different annotation layers into text corpora, and as such is one important step towards enabling corpus-pragmatic research to become more cumulative in future (see further Rissanen and Tyrkkö 2013, Hardie 2014). Thus, while the first steps have already been taken, more remains to be done (see below).

As for corpus design and data quality (iii–iv), corpora used in the analysis of specialised languages typically range from small to moderately large.[18] The advantage of these kinds of corpora is the careful text selection following a structured sampling frame, which reduces the amount of "noise" in the findings and makes them easier to interpret. Systematic text selection also makes it easier to tackle pragmatically-oriented research questions and track them across centuries (see e.g. **Hiltunen**, **this volume**, **Walker and Kytö**, **this volume**). Annotating these corpora is also more straightforward than annotating mega-corpora, as manual adjustments can be made.[19] More recently, in the wake of the recent "data explosion" in linguistics (see e.g. Liberman 2014, Hiltunen et al. 2017, Kortmann 2018), historical medical discourse has also become accessible to big-data analyses using methods of computational linguistics. Through applying advanced statistical methods to large

---

17.   For example, the LMEMT corpus was released in two parallel versions: a TEI-XML version and an unannotated version (Hiltunen and Tyrkkö 2019).

18.   For example, the three corpora representing English medical writing since the late medieval period to the beginning of the nineteenth century – *Middle English Medical Texts* (MEMT; Taavitsainen et al. 2005), *Early Modern English Medical Texts* (EMEMT; Taavitsainen et al. 2010), and *Late Modern English Medical Texts* (LMEMT; Taavitsainen et al. 2019) – comprise in total roughly four million words of running text. In other words, in terms of size they are small compared to mega-corpora or "big data" archives, like the 400-million-word *Corpus of Historical American English* (COHA; Davies 2012), but on a par with the ARCHER corpus (*A Representative Corpus of Historical English Registers*; Biber et al. 1993) and most second-generation English historical corpora used in genre and register analyses.

19.   For example, the EMEMT corpus was released in two versions, the standard version and a spelling-normalised version, created by manually training the VARD (Variant Detector) spelling normalisation software (Lehto et al. 2011), which greatly reduces the number of errors in subsequent automated analyses, like collocate identification and part-of-speech tagging (e.g. Hiltunen and Tyrkkö 2013).

datasets, it is possible to describe rhetorical and discursive changes taking place over a long time period, for example the gradual waning of scholastic argumentation and the introduction and refinement of experimental and laboratory-based discourse (see **Schneider**, **this volume**). A prime example of this trend is the EEBO (*Early English Books Online*) corpus (**McEnery and Baker**, **this volume**; see also Pumfrey et al. 2012), which provides unprecedented opportunities for data-driven study of cultural and linguistic change. However, the EEBO corpus and other similar archives present a number of well-known problems for analysts, including erroneous metadata and varying quality of text due to OCR (optical character recognition) errors. It is also not straightforward to identify "medical books" in large, general databases, which means that this method might be more suitable for studying the role of medicine in the society at large, rather than developments that are specific to medical discourse. Alternatively, it is possible to identify medical books with the help of complex search terms and manual verification, which is often necessary to tackle specific research tasks (see also **Grego**, **this volume**). However, this slows down the process of automated data analysis and cuts down the amount of data available for investigation. Yet published medical books tell only part of the story, and to gain insight into patients' views, it is necessary to turn to other, more ephemeral genres, like correspondence (see **Plastina**, **this volume**).

One of the core issues of corpus-pragmatic research is the question of interpreting the findings from different perspectives (v–vi). Techniques of quantitative corpus linguistics (e.g. frequency analysis, collocation analysis, variationist analysis) – and Digital Humanities (e.g. Topic Modelling, Kernel Density Estimation, conceptual maps) – provide powerful tools for identifying patterns of change in large datasets and drawing generalisations about the entire register. Yet insights from computational analyses are persuasive only if they can be meaningfully placed in the context of earlier ideas and claims. This process often relies crucially on substantive expertise in areas like medical and cultural history (not only of the British Isles), as well as history of earlier varieties of English. And further: the transmission of ideas through translation accommodated the contents into the vernacular for a wider and more heterogeneous audience. This phase in the history of English as a specialised language of medicine marked a new phase towards disseminating knowledge wider in society and paved the way to later health information in mass media (see **Rovelli**, **this volume**).

It is also clear that the variation between individual discourse contexts and local text functions is considerable, and to obtain reliable interpretations, it is often necessary to turn to methods of qualitative discourse analysis, such as genre analysis, linguistic narrative analysis and Critical Discourse Analysis (see **Whitt**; **Plastina**; **Ollikainen**, **this volume**). Where qualitative questions are the point of departure, corpus evidence can provide added systematicity and exactness to

analyses. Different methods may also lead to similar or partially similar outcomes; for example, discourse topics can be identified through close reading (**Grego**, **this volume**), using Topic Modelling (**Schneider; Menzel, this volume**), or with the help of collocation analysis (**Lehto, this volume**), and further work is needed to establish what works best in different situations. All this underlines the importance of considering the data from multiple perspectives.

This volume aims to give an incentive for further explorations into the history of the special language of medicine. Researchers today are in the fortunate position to have several corpora and large digital databases easily available for developing new multidisciplinary approaches by combining old and new methods. The common goal of the present collection of articles is to find answers to pertinent research questions by solid empirical evidence collected by corpus linguistic methods assessed in multilayered sociohistorical contexts. Progress has been rapid.

Corpus pragmatics is a versatile field. The most important signposts for the future point towards two directions. First, corpus annotation is an area of rapid advancement with a great potential. Among the desiderata for future research is the need for more pragmatically annotated corpora as well as studies that demonstrate their use by reproducible analyses. By using annotated corpora, it is possible to achieve the set goals more efficiently with better precision. However, more work is required in devising the systems for automatic retrieval, but the results can profit the research community at large.

Another major desideratum is the integration of advanced techniques and innovations developed by Digital Humanities more firmly into corpus linguistics and pragmatics (cf. Mehl 2021). Multi- or interdisciplinary approaches make use of different methodologies in their respective fields, breaking new ground and providing opportunities for innovation. For example, the crossroads of cultural history and medicine abound in uncharted topics that can be approached with the help of appropriate data-mining methods. Chapters in this volume give examples of the ways to proceed and provide challenges for exciting research.

We conclude this chapter by quoting The Royal Society motto from 1663, *Nullius in Verba* ('on the word of no one'). It expresses commitment to establishing scientific truth through empirical science rather than by quoting authorities, and it applies to the present research paradigm, too.

References

Archer, Dawn, Jonathan Culpeper, and Matthew Davies. 2008. "Pragmatic Annotation." In *Corpus Linguistics. An International Handbook*. Volume 1, ed. by Anke Lüdeling, and Merja Kytö, 613–640. Berlin: Mouton de Gruyter.

Atkinson, Dwight. 1999. *Scientific Discourse in Sociohistorical Context. The Philosophical Transactions of the Royal Society of London, 1675–1975*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Atkinson, Dwight, and Ellen Valle. 2012. "Corpus Analysis of Scientific and Medical Writing Across Time." In *The Encyclopedia of Applied Linguistics*, ed. by Carol A. Chapelle. Chichester : Blackwell Publishing. https://doi.org/10.1002/9781405198431.wbeal0251

Bazerman, Charles. 1988. *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison, WI: University of Wisconsin Press.

Biber, Douglas, and Susan Conrad. 2019. *Register, Genre, and Style*. Cambridge: Cambridge University Press. 2nd edition. https://doi.org/10.1017/9781108686136

Biber, Douglas, and Edward Finegan. 1994. "Intra-Textual Variation within Medical Research Articles." In *Corpus-Based Research into Language. In Honour of Jan Aarts*, ed. by Nelleke Oostdijk, and Pieter de Haan, 201–221. Amsterdam: Rodopi.

Biber, Douglas, Edward Finegan, and Dwight Atkinson. 1993. "ARCHER and Its Challenges: Compiling and Exploring a Representative Corpus of Historical English Registers." In *Creating and Using English Language Corpora*, ed. by Udo Fries, Gunnel Tottie, and Peter Schneider, 1–13. Amsterdam: Rodopi

Blommaert, Jan. 2005. *Discourse: A Critical Introduction*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511610295

Brown, Michael. 2011. *Performing Medicine: Medical Culture and Identity in Provincial England, c. 1760–1850*. Manchester: Manchester University Press.

Bynum, W. F. 1994. *Science and the Practice of Medicine in the Nineteenth Century*. Cambridge: Cambridge University Press.

Clancy, Brian, and Anne O'Keeffe. 2015. "Pragmatics." In *The Cambridge Handbook of Corpus Linguistics*, ed. by Douglas Biber, and Randi Reppen, 235–51. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139764377.014

Conrad, Susan. 2015. "Register variation." In *The Cambridge Handbook of English Corpus Linguistics*, ed. by Douglas Biber and Randi Reppen, 309–329. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139764377.018

Davies, Mark. 2012. "Expanding Horizons in Historical Linguistics with the 400-Million Word *Corpus of Historical American English*." *Corpora* 7 (2): 122–157. https://doi.org/10.3366/cor.2012.0024

Dorgeloh, Heidrun. 2016. "The Interrelationship of Register and Genre in Medical Discourse." In *Variational Text Linguistics: Revisiting Register in English*, ed. by Christoph Schubert and Christina Sanchez-Stockhammer, 43–65. Berlin/New York: Mouton De Gruyter. https://doi.org/10.1515/9783110443554-004

Ferguson, Charles A. 1994. "Dialect, Register and Genre. Working Assumptions about Conventionalization." In *Sociolinguistic Perspectives on Register*, ed. by Douglas Biber and Edward Finegan, 15–30. Cambridge: Cambridge University Press.

Flanagan, Joseph. 2017. "Reproducible Research: Strategies, Tools, and Workflows." In *Big and Rich Data in English Corpus Linguistics: Methods and Explorations*. Studies in Variation, Contacts and Change in English 19, ed. by Turo Hiltunen, Tanja Säily, and Joe McVeigh. https://varieng.helsinki.fi/series/volumes/19/flanagan/index.html

Geertz, Clifford. 1973. "Thick Description: Toward an Interpretive Theory of Culture." In *The Interpretation of Cultures: Selected Essays*, 3–30. New York: Basic Books.

Gotti, Maurizio. 2001. "The Experimental Essay in Early Modern English." *European Journal of English Studies* 5 (2): 221–239. https://doi.org/10.1076/ejes.5.2.221.7307

Gotti, Maurizio, and Francoise Salager-Meyer (eds). 2006. *Advances in Medical Discourse Analysis: Oral and Written Contexts*. Bern: Peter Lang. https://doi.org/10.3726/978-3-0351-0396-0

Gotti, Maurizio, Stefania Maci, and Michele Sala (eds). 2015. *Insights into Medical Communication*. Bern: Peter Lang. https://doi.org/10.3726/978-3-0351-0844-6

Gray, Bethany Ekle, Douglas Biber, and Turo Hiltunen. 2011. "The Expression of Stance in Early (1665–1712) Publications of the *Philosophical Transactions* and Other Contemporary Medical Prose: Innovations in a Pioneering Discourse." In *Medical Writing in Early Modern English*, ed. by Irma Taavitsainen, and Päivi Pahta, 221–47. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511921193.013

Hardie, Andrew. 2014. "Modest XML for Corpora: Not a Standard, but a Suggestion." *ICAME Journal* 38: 73–103. https://doi.org/10.2478/icame-2014-0004

Hiltunen, Turo. 2010. "*Philosophical Transactions*." In *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen, and Päivi Pahta, 127–31. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.160.04cat6

Hiltunen, Turo. 2021. Intensification in Eighteenth Century Medical Writing. *Journal of English Linguistics*, 49 (1), 90–113. https://doi.org/10.1177/0075424220982649

Hiltunen, Turo, and Jukka Tyrkkö. 2013. "Tagging *Early Modern English Medical Texts*." Paper presented at the *Corpus Analysis with Noise in the Signal (CANS)* workshop, Lancaster University, 22nd July 2013. http://ucrel.lancs.ac.uk/cans2013/

Hiltunen, Turo, and Jukka Tyrkkö. 2019. "Manual to the LMEMT Corpus." In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221.12hil

Hiltunen, Turo, Tanja Säily, and Joe McVeigh. 2017. "How to Turn Linguistic Data into Evidence." In *Big and Rich Data in English Corpus Linguistics: Methods and Explorations*. Studies in Variation, Contacts and Change in English 19, ed. by Turo Hiltunen, Tanja Säily, and Joe McVeigh. https://varieng.helsinki.fi/series/volumes/19/introduction.html

Honkapohja, Alpo. 2017. *Alchemy, Medicine, and Commercial Book Production*. Turnhout: Brepols. https://doi.org/10.1484/M.TT-EB.5.111529

Jacobs, Andreas, and Andreas H. Jucker. 1995. "The Historical Perspective in Pragmatics." In *Historical Pragmatics: Pragmatic Developments in the History of English*, ed. by Andreas H. Jucker, 3–33. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/pbns.35.04jac

Jucker, Andreas H. 2008. "Historical Pragmatics." *Language and Linguistic Compass* 2 (5): 894–906. https://doi.org/10.1111/j.1749-818X.2008.00087.x

Jucker, Andreas H. 2018. "Corpus Pragmatics" In *Methods in Pragmatics*, ed. by Andreas H. Jucker, Klaus P. Schneider, and Wolfram Bublitz, 455–66. Mouton de Gruyter: Berlin. https://doi.org/10.1515/9783110424928-018

Jucker, Andreas H., Daniel Schreier, and Marianne Hundt. 2009. "Corpus Linguistics, Pragmatics and Discourse." In *Corpora: Pragmatics and Discourse, Papers from the 29th International Conference on English Language Research on Computerized Corpora (ICAME 29)*, ed. by Andreas H. Jucker, Daniel Schreier, and Marianne Hundt, 3–7. Amsterdam: Rodopi.

Jucker, Andreas H., and Irma Taavitsainen. 2014. "Diachronic Corpus Pragmatics: Intersections and Interactions." In *Diachronic Corpus Pragmatics*, ed. by Irma Taavitsainen, Andreas H. Jucker, and Jukka Tuominen, 3–26. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/pbns.243.03juc

Jurin, James. 1724. *An Account of The Success of Inoculating The Small Pox in Great Britain*. London: J. Peele.

Kohnen, Thomas. 2009. "Historical Corpus Pragmatics: Focus on Speech Acts and Texts." In *Corpora: Pragmatics and Discourse. Papers from the 29th International Conference on English Language Research on Computerized Corpora (ICAME29)*, ed. by Andreas H. Jucker, Daniel Schreier, and Marianne Hundt, 13–36. Amsterdam: Rodopi.

Kortmann, Bernd. 2018. "Reflecting on the Quantitative Turn in Linguistics." Online lecture. University of Freiburg. https://bit.ly/2VlgFCL. (Accessed 9 February 2021.)

Labov, William. 1972. "Some Principles of Linguistic Methodology." *Language in Society* 1 (1): 97–120. https://doi.org/10.1017/S0047404500006576

Lehto, Anu, Alistair Baron, Maura Ratia, and Paul Rayson. 2011. "Improving the Precision of Corpus Methods: The Standardized Version of *Early Modern English Medical Texts*." In *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen and Päivi Pahta, 279–290. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.160.11leh

Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511813313

Liberman, Mark. 2014. "How Big Data is Changing How We Study Languages." *The Guardian*, 7 April 2014. https://www.theguardian.com/education/2014/may/07/what-big-data-tells-about-language

Marchi, Anna, and Charlotte Taylor. 2018. "Introduction: Partiality and Reflexivity." In *Corpus Approaches to Discourse. A Critical Review*, ed. by Charlotte Taylor, and Anna Marchi, 1–16. London: Routledge. https://doi.org/10.4324/9781315179346-1

Martin, Gillian. 2014. "Pragmatics and Medical Discourse." *Pragmatics of Discourse*, ed. by Klaus P. Schneider, and Anne Barron, 491–524. Berlin, München, Boston: Mouton De Gruyter. https://doi.org/10.1515/9783110214406-019

Mehl, Seth. 2021. "Why Linguists Should Care about Digital Humanities (and Epidemiology)." *Journal of English Linguistics*, 49 (3), 331–337. https://doi.org/10.1177/00754242211019072

Mey, Jacob L. 2001. *Pragmatics: An Introduction*. Second edition. Oxford: Blackwell.

Millar, Neil, Brian Budgell, and Keith Fuller. 2013. "'Use The Active Voice Whenever Possible': The Impact of Style Guidelines in Medical Journals." *Applied Linguistics* 34 (4): 393–414. https://doi.org/10.1093/applin/ams059

Moessner, Lilo. 2008. "Variation and Change in the Writings of 17th-Century Scientists." In *English Historical Linguistics 2006. Volume II: Lexical and Semantic Change*, ed. by Richard Dury, Maurizio Gotti and Maria Dossena, 75–92. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/cilt.296.07moe

Nevalainen, Terttu. 1999. "Making the Best Use of "Bad" Data: Evidence for Sociolinguistic Variation in Early Modern English." *Neuphilologische Mitteilungen* 100 (4): 499–533.

Nwogu, Kevin Ngozi. 1997. "The Medical Research Paper: Structure and Functions." *English for Specific Purposes* 16 (2): 119–138. https://doi.org/10.1016/S0889-4906(97)85388-4

Ordóñez-López, Pilar and Nuria Edo-Marzá. 2016. "Medical Discourse: Building Bridges between Medicine and Society." *Medical Discourse in Professional, Academic and Popular Settings*, ed. by Pilar Ordóñez-López, and Nuria Edo-Marzá, 1–8. Bristol: Multilingual Matters. https://doi.org/10.21832/9781783096268-002

Pahta, Päivi. 1998. *Medieval Embryology in the Vernacular: The Case of De spermate* [Mémoires de la Société Néophilologique de Helsinki 53]. Helsinki: Société Néophilologique.

Pahta, Päivi and Irma Taavitsainen. 2004. "Vernacularisation of scientific and medical writing in its sociohistorical context." In *Medical and Scientific Writing in Late Medieval English*, ed. by Irma Taavitsainen and Päivi Pahta, 1–22. Cambridge: Cambridge University Press.

Piqué-Angordans, Jordi, and Santiago Posteguillo. 2006. "Medical Discourse and Academic Genres." In *Encyclopedia of Language & Linguistics*, ed. by Keith Brown, 649–657. Oxford: Elsevier. https://doi.org/10.1016/B0-08-044854-2/02350-6

Pumfrey, Stephen, Paul Rayson, and John Mariani. 2012. "Experiments in 17th Century English: Manual versus Automatic Conceptual History". *Literary and Linguistic Computing* 27 (4): 395–408. https://doi.org/10.1093/llc/fqs017

Redden, Elizabeth. 2020. "Rush to Publish Risks Undermining COVID-19 Research." *Inside Higher Education*, 8 June 2020. https://www.insidehighered.com/news/2020/06/08/fast-pace-scientific-publishing-covid-comes-problems

Rissanen, Matti, and Jukka Tyrkkö. 2013. "*The Helsinki Corpus of English Texts* (HC)." In *Principles and Practices for the Digital Editing and Annotation of Diachronic Data*. Studies in Variation, Contacts and Change in English 12, ed. by Anneli Meurman-Solin, and Jukka Tyrkkö. Helsinki: VARIENG. https://varieng.helsinki.fi/series/volumes/14/rissanen_tyrkko/

Rockwell, Geoffrey. 2013. "Inclusion in Digital Humanities." In *Defining Digital Humanities: A Reader*, ed. by Melissa Terras, Julianne Nyhan and Edward Vanhouette. Farnham: Ashgate.

Romero-Trillo, Jesús. 2008. *Pragmatics and Corpus Linguistics: A Mutualistic Entente*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110199024

Romero-Trillo, Jesús. 2017. "Editorial: Corpus Pragmatics." *Corpus Pragmatics* 1: 1–2. https://doi.org/10.1007/s41701-017-0005-z

Rühlemann, Christoph, and Karin Aijmer. 2015. "Corpus Pragmatics: Laying the Foundations." In *Corpus Pragmatics: A Handbook*, ed. by Karin Aijmer and Christoph Rühlemann, 1–26. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139057493.001

Salager-Meyer, Françoise, María Ángeles Alcaraz Ariza, and Maryelis Pabón Berbesí. 2007. "Collegiality, Critique and the Construction of Scientific Argumentation in Medical Book Reviews: A Diachronic Approach." *Journal of Pragmatics* 39: 1758–1774. https://doi.org/10.1016/j.pragma.2006.06.003

Smith, Nicholas, Sebastian Hoffmann, and Paul Rayson. 2008. "Corpus Tools and Methods, Today and Tomorrow: Incorporating Linguists' Manual Annotations." *Literary and Linguistic Computing* 23 (2): 163–180. https://doi.org/10.1093/llc/fqn004

Swales, John. 2000. "Language for Specific Purposes." *Annual Review of Applied Linguistics* 20: 59–76. https://doi.org/10.1017/S0267190500200044

Taavitsainen, Irma. 2000. Scientific Language and Standardisation 1375–1550. In *The Development of Standard English: Theories, Descriptions, Conflicts*, ed. by Laura Wright, 131–154. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511551758.009

Taavitsainen, Irma. 2018. "Scholastic Genre Scripts in English Medical Writing 1375–1800." In *Diachronic Corpora, Genre, and Language Change*, ed. by Richard J. Whitt, 95–115. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/scl.85.05taa

Taavitsainen, Irma. 2019. "Polite Society Language Practices: Letters to the Editor in The Gentleman's Magazine." In Irma Taavitsainen and Turo Hiltunen (eds). 2019. *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, 129–144. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221.07taa

Taavitsainen, Irma. 2020. "A Medical Debate of 'Heated Pamphleteering' in the Early Eighteenth Century." In *Manners, Norms and Transgressions in the History of English: Literary and Linguistic Approaches*, ed. by Andreas H. Jucker, and Irma Taavitsainen, 141–163. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/pbns.312.07taa

Taavitsainen, Irma. 2021. "Scientific book reviews in the seventeenth century: Early stages of the genre." In *The Role of Context in the Production and Reception of Historical News Discourse*, ed. by Nicholas Brownlees. 67–86. Bern: Peter Lang.

Taavitsainen, Irma and Susan Fitzmaurice. 2007. "What It Is and How to Do It." In *Methods in Historical Pragmatics*, ed. by Susan M. Fitzmaurice and Irma Taavitsainen, 11–36. Berlin: Walter de Gruyter. https://doi.org/10.1515/9783110197822.11

Taavitsainen, Irma, and Turo Hiltunen (eds). 2019. *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221

Taavitsainen, Irma and Turo Hiltunen. Forthcoming. "Medical and Scientific Writing." In *New Cambridge History of the English Language. Volume II: Documentation, sources of data and modelling*, ed. by Merja Kytö and Erik Smitterberg. Cambridge: Cambridge University Press

Taavitsainen, Irma, Peter Murray Jones, and Turo Hiltunen. 2019. "Sociohistorical and Cultural Context of *Late Modern English Medical Texts*." In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Turo Hiltunen, and Irma Taavitsainen, 17–30. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221.02taa

Taavitsainen, Irma, and Päivi Pahta. 2010. "Scientific Discourse." In *Handbook of Historical Pragmatics*, ed. by Andreas H. Jucker, and Irma Taavitsainen, 3–32. Berlin; New York: Mouton de Gruyter. https://doi.org/10.1515/9783110214284.1.3

Taavitsainen, Irma, and Andreas H. Jucker. 2015. "Twenty Years of Historical Pragmatics: Origins, Developments and Changing Thought Styles." *Journal of Historical Pragmatics* 16 (1): 1–24. https://doi.org/10.1075/jhp.16.1.01taa

Taavitsainen, Irma, and Päivi Pahta (eds). 2010. *Early Modern English Medical Texts: Corpus Description and Studies*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.160

Taavitsainen, Irma, and Päivi Pahta (eds). 2011. *Medical Writing in Early Modern English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511921193

Taavitsainen, Irma, Turo Hiltunen, Anu Lehto, Ville Marttila, Päivi Pahta, Maura Ratia, Carla Suhr and Jukka Tyrkkö. 2019. *Late Modern English Medical Texts: The Corpus*. In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221

Taavitsainen, Irma, Päivi Pahta, Turo Hiltunen, Martti Mäkinen, Ville Marttila, Maura Ratia, Carla Suhr, and Jukka Tyrkkö. 2010. *Early Modern English Medical Texts: Corpus*. In *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen, and Päivi Pahta. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.160

Taavitsainen, Irma, Päivi Pahta, and Martti Mäkinen. 2005. *Middle English Medical Texts*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.131

Taavitsainen, Irma, Jeremy J. Smith, Carla Suhr, and Turo Hiltunen. (Forthcoming). Genre in Early Modern Medical English: Sociocultural Contexts of Production and Use, 1500–1820. (Studies in English Language). Cambridge: Cambridge University Press.

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/scl.6

Traugott, Elizabeth. 2008. "The State of English Language Studies: A Linguistic Perspective." In *English Now: Selected Papers from the 20th IAUPE Conference in Lund 2007*. Lund Studies in English 112, ed. by Marianne Thormählen, 199–225. Lund, Sweden: Lund University.

Verschueren, Jef. 2013. "Continental European Perspective View." In *The Oxford Handbook of Pragmatics*, ed. by Yan Huang. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199697960.013.9

Voigts, Linda E. 1984. "Medical Prose." In *Middle English Prose: A Critical Guide to Major Authors and Genres*, ed. by A. S. G. Edwards, 315–335. New Brunswick: Rutgers.

Voigts, Linda E. 1989. "Scientific and Medical Books." In *Book Production and Publishing in Britain, 1375–1475*, ed. by Jeremy Griffths, and Derek Pearsall, 345–402. Cambridge: Cambridge University Press.

Wear, Andrew. 2000. Knowledge and Practice in English Medicine, 1550–1680. Cambridge: Cambridge University Press.

Webster, Charles. 1975. *The Great Instauration: Science, Medicine and Reform 1626–1660*. London: Duckworth.

Tracing discursive changes

"A geography of names"

A genre analysis of nationality-driven names for venereal disease in seventeenth-century England

Tony McEnery and Helen Baker
University of Lancaster

We investigate discourses surrounding venereal disease in a wide body of seventeenth-century texts in the one-billion-word *Early English Books Online* (EEBO) corpus. By combining quantitative methods with close reading of texts within EEBO, we explore whether perceptions of sufferers and responses to the illness itself shifted as the century progressed. In order to uncover the kinds of written works in which references to venereal disease appear, we undertake a genre analysis with the help of genre categories added to the corpus. The analysis shows that at the beginning of the century, references to these diseases were most likely to appear within historical works, while the end of the century witnesses an increase in references to venereal disease in scientific and medical texts.

Keywords: genre, onomastics, terminology, venereal disease, nationality, Early English Books Online (EEBO)

1.   Introduction

In this chapter, we investigate how venereal disease was written about in seventeenth-century England. After establishing what words we may want to search for to explore venereal disease in the period, we investigate the occurrence of "nationality"-driven terms for venereal disease and show that English writers were most likely to describe the disease as originating from the French. We move on to illustrate the potential of a genre classification system which has been recently applied to the EEBO-TCP corpus (Murphy, 2019) to allow us to discover what kinds of written works most commonly mentioned these nationality-driven names and in what ways. Alongside discussing how symptoms were described and what kinds of treatments were offered for the diseases, we offer evidence of how perceptions

surrounding sufferers and responses to venereal disease itself shifted as the century progressed. To begin with, however, we will briefly consider the link of place and disease, the data used in this study and the approach to genre taken in that data.

1.1   Disease and place

The attribution of a disease to a geographic location was particularly meaningful at the time this chapter was written. In mid-March 2020, US President Donald Trump defended his references to the SAR-CoV-2 virus responsible for the global coronavirus pandemic as the "China virus" and the "Chinese virus": "Cause it comes from China. It's not racist at all, no, not at all." A week or so later, he announced he would no longer be using these terms but insisted he had no regrets about previously employing them, reiterating that "everyone knows [the virus] came out of China" and that it was customary to reference diseases in terms of their country of origin.[1]

Trump cited Lyme Disease and Ebola, named after a town and river respectively, as examples of this custom but another high-profile example is the so-called Spanish flu of 1918–20, inaccurately thought to have originated in Spain or to have affected that country's inhabitants particularly severely. These names matter. As Keiji Fukuda, Assistant Director-General for Health Security of the World Health Organization, has commented: "We've seen certain disease names provoke a backlash against members of particular religious or ethnic communities, create unjustified barriers to travel, commerce and trade, and trigger needless slaughtering of food animals. This can have serious consequences for peoples' lives and livelihoods."[2] Accordingly, the WHO selected the name COVID-19 for the coronavirus disease which was first identified in Wuhan in December 2019. Likewise, later variants of the disease, while initially linked to placenames, were ultimately named without reference to placenames precisely because of the stigma that a link between place and disease can bring to a location and its inhabitants.

Yet in the seventeenth century, the attribution of venereal disease to particular countries carried twice the stigma because as well as the health concerns related to the disease, this disease was known to be transmitted via sexual activity which, in itself, carried a burden of shame and secrecy. Venereal disease cast a burden on the rich and poor alike – it did not differentiate between social strata. Any person could unknowingly become infected, providing they had engaged in sexual relations with an infected person. The spread of the disease was facilitated, particularly

---

1.   See Vazquez and Klein (2020) and Vazquez (2020).

2.   WHO, see https://www.who.int/news/item/08-05-2015-who-issues-best-practices-for-naming-new-human-infectious-diseases.

in the early years after contraction, because it was possible to disguise symptoms with reasonable success.

Unsurprisingly in this context, venereal disease was a prevalent problem in early modern England. It remains uncertain when the first cases of the disease appeared in England but, as early as 1161, an ordinance of Henry II, legalising the Southwark brothels, referred to "the perilous disease of burning" (Linane 2003: 27). The founding of the Lock Hospital, ostensibly to treat lepers, in Southwark by Edward II in 1321 suggests that gonorrhoea and syphilis may have posed a threat to the health of England's population by this time. It is uncertain what proportion of the early-modern population contracted sexually transmitted infections throughout their lifetimes, but historians believe that figures are high. Siena (2001: 205) argues that historians may have previously underestimated the prevalence of venereal disease. He notes that around 20 to 30 per cent of hospital patients in London were admitted with the disease. Szreter (2017) has recently attempted to quantify the incidence of syphilis in mid-1770s Chester, estimating that around 8 per cent of the city's population had been infected by the time they reached the age of 35.

Early moderners did not successfully distinguish between different types of sexually transmitted diseases. For instance, the symptoms of gonorrhoea and syphilis were viewed as manifestations of different stages of the same illness.[3] Indeed, as Qualtiere and Slights (2003: 6) have noted, the transformative character of the pox was such that some early modern works, for instance Peter Lowe's *Easie, certaintie, and perfect method, to cure and preuent the Spanish sicknes* (1569), claimed that the disease could lead to leprosy. We will follow Szreter (2017: 210) by using the more neutral phrase *venereal disease*.

2. Material and methods

2.1 Corpus

Our study uses EEBO-TCP, a corpus which consists of 44,422 texts – for the seventeenth century, the word count approaches one billion words.[4] Table 1 shows how the corpus is distributed in terms of number of words throughout the decades. The spread is fairly even in the first four decades, but greater fluctuation is evident in the rest of the century; a modest jump in numbers of words in the 1640s is followed by a much larger increase in the 1650s when the corpus peaks in size. EEBO-TCP

---

3. See Churchill (2005: 7) and Siena (1998: 556).

4. For more information about EEBO-TCP see: https://quod.lib.umich.edu/e/eebogroup/.

is drawn from EEBO (Early English Books Online), a collection of over 125,000 works published in England, Ireland, Scotland, Wales and British North America between the years 1473 and 1700. Because EEBO-TCP was compiled by manually keying in letters, the accuracy level of the digital text is very high, at 99.995 per cent.[5] In terms of the number of texts, EEBO-TCP represents around 35 per cent of EEBO. Texts were not randomly selected for inclusion in EEBO-TCP; rather, curators "aimed to key as many different works – as much different *text* – as possible". Priority was given to first editions and English-language works while serials, texts with limited textual content and those which were poorly preserved were avoided.[6] Hamlin (2014: 18) has noted that EEBO-TCP is "broadly representative of EEBO's full coverage".

Table 1. Corpus size per decade in the seventeenth-century EEBO-TCP data

Span	No. of words
1600–1609	57,272,438
1610–1619	61,922,837
1620–1629	55,748,866
1630–1639	63,496,401
1640–1649	87,480,996
1650–1659	168,912,438
1660–1669	111,998,646
1670–1679	118,167,747
1680–1689	142,071,417
1690–1699	128,494,904
1600–1699 (whole century)	996,472,953

EEBO-TCP was accessed via CQPweb, Lancaster University's online corpus analysis system, which allows researchers to search and manipulate the data in order to gain insight into discourse patterns. In this study, we used CQPweb to quickly compile frequency data surrounding language relating to venereal disease and to generate collocates. Collocates are terms which are more likely to appear alongside a word of interest than elsewhere in the corpus.[7]

---

5.   For a discussion of how the electronic text was generated, see https://textcreationpartnership. org/using-tcp-content/results-of-keying/.

6.   For a more detailed discussion of how texts were selected, see https://textcreationpartnership. org/tcp-texts/eebo-tcp-early-english-books-online/.

7.   For information relating to CQPweb, see https://cqpweb.lancs.ac.uk/ and Hardie (2012).

2.2   Genre

Up until very recently, the EEBO-TCP corpus provided users with metadata of a limited nature, for example indicating texts' authorship, date and place of publication (where known) but with no genre classifications. In previous investigations we have been able to identify when certain terms tended to appear in one particular genre by means of a qualitative analysis. For example, in our study of early modern vagrancy (McEnery and Baker, 2017a) we found that the term *rogue* increased in usage in play dialogue in the second half of the seventeenth century. However, such genre analysis has so far been somewhat impressionistic, with identification tending only to be successful when a node or collocate was overwhelmingly linked to a specific genre at a specific time to the extent that the association became self-evident in concordancing and close reading.

As a result of work undertaken on genre classification, significant progress has been made in constructing a workable genre framework for titles within the EEBO-TCP corpus which has enabled us to take a genre-based approach in this study. The categorisation framework is based upon the digital identification of genre, as indicated by the title of each piece of writing in the EEBO-TCP corpus. The scheme was created by Tony McEnery and further developed by Sean Murphy, who describes his rationale in Murphy (2019). As Murphy (ibid.) has explained, the term *genre* is understood to refer to "a category of texts grouped according to culturally, conventionally and consensually recognised criteria which change over time and which allow for division into sub-genres (e.g. letters, essays.)"[8] *Domain*, meanwhile, refers to broad subject fields which might encompass a number of genres; the five domains of EEBO-TCP are Literary (which contains five genres, Plays; Poetry; Verse & Song; Fiction; and General), Religious (containing five genres of Bible; Catholicism; Protestantism; Doctrine, Theology & Governance; and General), Administrative (split into the four genres of Royal; Parliamentary; Legal; and General), Instructional (with five genres of Philosophical; Science; Mathematics; Medicine; and General), and, lastly, Informational (with six genres of Biography; Colonial; Essay; Letters; Pamphlets; and General). Some genres are divided again into sub-genres. For instance, Plays breaks down into four sub-genres: Comedy; History; Tragedy; and Masque.[9]

---

8.   Also see explanations of the terms genre and text-type by Taavitsainen (2001a: 139–140) and Smitterberg and Kytö (2015: 118–119).

9.   The categories discussed come from McEnery's scheme as Murphy's has yet to be applied to the EEBO-TCP corpus. For this study we were able to access an early prototype of the corpus annotated with McEnery's classification. That prototype classification has been applied to 92.7% of the texts in the corpus. It is available on the Lancaster University CQPweb installation: https://cqpweb.lancs.ac.uk/.

The array and arrangement of the categories within the whole EBO-TCP corpus immediately tells us useful information about early modern English society. For instance, the abundance of Religious sub-genres indicates, quite rightly, how religious texts saturated early modern writing. As Smitterberg and Kytö (2015: 119) emphasise, genres develop and change through time, and this can be seen within the EEBO-TCP corpus itself. Masques, for instance, lost popularity in the latter half of the seventeenth century, which is reflected in the dwindling of the sub-genre Plays: Masques. Moreover, the ways in which authors or booksellers chose to label their work, often for marketing purposes, also underwent changes due to shifting literary fashions and tastes. This also has repercussions for genre identification. For instance, Orr (2011: 72) has shown how the popularity of the terms *novel* and *romance* on title pages fluctuated in the seventeenth and eighteenth centuries, whilst fictional works with labels such as *history* and *memoirs* relied on the contemporary reader to pick up other clues as to their lack of adherence to factual events. Authors' self-labelling by genre was important to readers. Sullivan (2007) has explained how lengthy early modern titles indicated genre to the reader and, in turn, guided their recognition and expectations of a work's content before they had even opened its pages.[10]

The use of genre analysis in historical research is not without controversy,[11] but historical linguistic research by scholars such as Taavitsainen (2001b), Ratia (2013) and Bergs (2004) has shown how genre identification and analysis can enrich and sharpen our understanding of an historical text and thereby illuminate its function and its author's intent. For historians, an assessment of genre can throw light upon the representation of a particular group or topic as that representation shifts over time. For instance, we may be able to uncover, at scale, patterns in the types of texts in which our words of interest appear. The addition of genre categorisation within the EEBO-TCP corpus offers researchers the opportunity to accomplish a deeper layer of analysis and, combined with the capabilities of the online concordancer, CQPweb, greater flexibility to pursue individual research questions and restrict research to specific genres and/or periods of time.[12]

---

10. See also Taavitsainen (2001a: 140) who explains that genres function as "'horizons of expectation' for readers to know what to expect and models of writing for authors".

11. See Cohen (1986).

12. See Murphy (2019).

3.   Results

3.1   What words were used to refer to venereal disease?

The first part of our analysis necessitated the compilation of a list of words which early modern writers used to refer to venereal disease. We determined these words by means of five stages of research: (i) reading a wide range of relevant works by historians; (ii) consulting the work of contemporary authors; (iii) searching digital libraries of key historical sources, such as British History Online and London Lives;[13] (iv) performing keyword searches in Lexicons of Early Modern English (a historical database of dictionaries, encyclopaedias and glossaries),[14] Green's Dictionary of Slang,[15] and the Historical Thesaurus of the Oxford English Dictionary (OED);[16] and, lastly, (v) utilising information garnered from the corpus itself, mostly in the exploration of collocates of words suggested in i-iv and by close reading of concordances of those words.

Figure 1 shows the results of this process. The list of names is divided into seven broad categories based on how they were derived. For example, the Symptoms category lists names for venereal disease which can be traced back to the characteristic physical manifestations or discomfort associated with the illness. There is some overlap between these categories. For instance, it could be argued that *pox* should be included in the Specific Disease category due to there being a number of named diseases which end in the suffix *pox*, but here it is included in the Symptoms category because the word *pocks* means pustules. As Figure 1 shows, many words in our list are constructed in a similar way, for instance by the insertion of a nationality adjective in front of the term *pox* or *disease* (e.g. French disease, Italian disease, American disease) or by the same adjective modifying a number of different nouns from the same semantic field (e.g. venereal disease, venereal distemper, venereal plague).

It is important to note that Figure 1 does not represent an exhaustive list. It is almost certain that there are more of these terms and their variants which we have missed. Nonetheless, we are confident that the list of words we have chosen to investigate represents, at the very least, a substantial and broad set of terms used to refer to venereal disease in the seventeenth century.

---

13. https://www.londonlives.org/ and https://www.british-history.ac.uk/

14. http://leme.library.utoronto.ca/ and Lancashire (2015).

15. https://greensdictofslang.com/ GDoS Online was launched in October 2016 and was made available for free in November 2018.

16. http://www.oed.com/thesaurus

Specific Disease	gonorrhoea, [gomorra, gomer, gomoria, gomorrea, gomory, gonorrhey, gonorrhaea, gonorrhea, gonorhea], gonorrhaean passion, gout, syphilis [siphylis, siphile, siphilis, syphylis], venereal syphilis
Sexual Activity	a crack,[16] Cupid's disease, lues venerea [venrea], the venereal, venereal bubo, venereal disease, venereal distemper, venereal lues, veneraean murrain, venereal plague, the venereous evil, venerous pox, venery, the whores pox
Nationality	French alamode, French bone-ache, French cannibal [French canniball], French chilblains [French chilbalines], French crown [French-crowne], French crust, French curse, the French disease, French distemper, French evil, French fever, French goods, French gout, French lues, French marbles, French measles, French morbus, French pestilence,[17] French plague, French pox [French pockes, French-pox], French razor, French scab, French sickness, lues gallica, malady of France, gallicus morbus, morbus gallicus, mal de Naples, morbus neapolitanus, the Naples disease, Naples pox, Neapolitan bone-ache, the Neapolitan disease, the Neopolitane, Neopolitanus, Hispanicus, Spanish disease, Spanish gout, Spanish scab [Spanish scabbe], Spanish pox, the Aleppo evil, the American disease, the Christian disease, the English pox, the disease of the Turk, the Frank disease, the Indian pox, Indicus, the Italian disease, the Polish disease
Symptoms	clap[18] [clappe], the pox, a gleet [gleat], the grand pox, the great pox, bube [bubo, the bubos], burning, the crinkomes[19] [crinkum(s), greencomes, grincam, grincom(s), grincome(s), griuncum(s), grinkcome(s), grinkham(s), grinkum(s)], gross verole, running of the reins,[20] scabbado, verola [verol]
Shame or Repulsion	affliction, the loathsom disease, the foul disease, the foul evil,[21] gore, grandgore[22] [grandgor, grantgore, grand gore, glengore, glengare], the loathed disease, the secret disease
Infectious	lues, pandora's box, pestilence, poore mortalls plague
Unknown	the black lion, a blow, flapdragon [flap dragon], nap

Figure 1. Terms of interest compiled using lexicography, background reading and a preliminary corpus analysis. Variant spellings are included in square brackets

17. *Crack* was also a nickname for a female prostitute.

18. *French pestilence* occurs in one late sixteenth-century EEBO-TCP text, see Marston (1598).

19. The old French term *clapoir* means a venereal sore.

20. Green believes that *crinkum* and its variant terms relate to "the sense of twisting pain that accompanies the disease" – hence its inclusion in the Symptoms category.

21. The phrase *running of the reins* refers to kidney discharge.

22. Siena (2001: 200) writes that patients suffering from venereal disease in early modern London were commonly termed *foul patients* and that they were admitted to *foul wards* but neither of these phrases are present in the EEBO-TCP corpus.

23. *Grandgore* was a Scottish term.

Relatively few words were identified in both the preliminary reading and initial corpus analysis stages. This indicates the importance of both stages but also suggests that some of the words arising from our reading of secondary sources may not have been as popularly used as historians believe. Phrases such as the *disease of the Turk* (Siena 1998: 572) and the *Polish disease* (Qualtiere and Slights 2003: 5) do not appear in the EEBO-TCP corpus at all. Spongberg (1997: 3) writes that syphilis was known as *grand pox* in the sixteenth century but there are only seven entries for this phrase in EEBO-TCP and the earliest of these, Cavendish (1649), was printed in the mid-seventeenth century.

Indeed, the textual reality behind the names included in this long list is that the vast majority of them appear very rarely in the EEBO-TCP corpus. Query searches for *Cupid's disease, veneraean murrain, venereal plague/distemper/syphilis, French crust* and *poore mortalls plague* achieved no hits at all. Other terms, which initially seemed more promising, such as *secret disease, loathsome disease* and *foul disease* were often revealed, when examples were subjected to close reading, either to refer to a different medical condition such as leprosy or were used figuratively. The terms *gout* and *pestilence* occur in very high numbers in the EEBO-TCP corpus but findings via two methods – reading a random sample of 100 concordances which mentioned each of these terms and studying their collocates – suggested that they were not frequently used to mean venereal disease specifically.[24] The only terms in our list which appeared moderately frequently over the course of the seventeenth century and almost always referred to venereal disease were *gonorrhoea* (941 hits, including its spelling variants), *running of the reins* (519 hits), *lues venerea* (402 hits), *French pox* (1233 hits) and *French disease* (537 hits).

We initially attempted to use the term POX as a starting point for our corpus analysis as this was by far the most frequent term used to refer to venereal disease we had encountered.[25] However, it quickly became apparent that POX is a problematic term because, as its collocates clearly showed, it was frequently used to refer to smallpox as well as venereal disease. Given that the term *French pox* appeared to be the most commonly used alternative to *the pox*, we decided to focus on this term, alongside other terms that attributed venereal disease to France. Indeed, the size of the "nationality" category in Figure 1 far outweighed any other, suggesting that a sizeable proportion of names for venereal disease included some reference to other countries.

---

24. The terms *French-pox, gonorrhoea* and *French* collocate with *gout* but this occurs in texts which list these diseases as being separate illnesses.

25. We included the variants *pockes* and *pocks* in the search query.

A revealing text by B.E. (1699), comparing naming conventions used for gypsies to those applied to venereal disease, highlighted the tendency of nations to blame their neighbours for being the source of the disease:

> For the Gypsies, they and the Foul Disease have alike the Fate to run through a Geography of Names, and to be made free of as many Countries, as almost there are Languages to call them Names in; for as the French call the Pox, the Italian Disease, they again give it to the Spaniards, as these to the French; so the French call the Gypsies Boemie, or Bohemians, belike, because they made their first Appearance in Bohemia of any Part of Europe; the Italians Name them Zingari or Saracen's, the Spaniards Itanos as we Egyptians; whether it be, that the Italians give them the Turks, as the Spaniards give them the Moors, as being both the next Neighbours to each; I take not upon me to Determine, only it may be observed, betwixt the Complement of either kind, the Odds is no greater than this, of giving a Nation a Clap, or of laying a brood of Bastards at it's Door.
>
> (B.E., *A New Dictionary of the Canting Crew, 1699*)

Qualtiere and Slights (2003: 5) have described how seventeenth-century writers struggled to find an appropriate technical name for venereal disease. Physicians throughout Europe, wishing to sidestep national responsibility for such a humiliating sickness, resorted to linking the name of other nations to the disease. As a result, in referencing venereal disease, a number of country names were used to premodify the term *disease* – Siena (1998: 572) specifically mentions the French Disease, the Neapolitan Disease, and the Disease of the Turk. Qualtiere and Slights (ibid.) assert that the Spanish, Polish, or Neapolitan Disease were frequently used but claim that by the mid-sixteenth century, the name Morbus Gallicus, or the French pox, was the most common way of referring to venereal disease. Spongberg (1997) notes the tendency for nations to blame their closest neighbours for the spread of venereal disease, noting that the Russians named it *Polish sickness*, the Poles referred to the *German sickness*, the Italians and French blamed one another, and the Chinese blamed the Japanese.

Donald Trump's comments reproduced in 1.1 suggest that the tendency to attribute disease to foreigners should not be regarded as a purely xenophobic attack upon the inhabitants of neighbouring countries. His comments rested upon a claim regarding the perceived historical origins of that disease. Might the same be true of writers in our corpus? It is unknown when the first cases of syphilis appeared in England but an influential explanation, sometimes known as the Columbian Exchange theory, posits that the disease was introduced to Europe by Spanish soldiers who had accompanied Christopher Columbus on his first trip to the New World. The captain of the Pinta, Martin Alonso Pinzón, died in March 1493, possibly suffering from the disease himself. Many of Columbus's crewmen went on to

serve Charles VIII of France during his invasion of Italy in 1494 where it is likely, according to scholars such as Kohn (2008: 130–131), that syphilis was spread between both armies, particularly once the French had reached Naples the following year. These troops contained mercenaries of Spanish, German, Swiss, English, Hungarian and Polish nationality who probably spread the disease to their home countries on their return. It may have been the case that the illness passed around at Naples was a more virulent strain of syphilis. Nevertheless, there is overwhelming evidence for its rapid distribution throughout Europe in the years immediately following the French invasion of Italy. In 1495, the Emperor Maximilian issued a proclamation that warned against the "evil pox". In 1496, every person infected with syphilis was ordered to leave Paris within 24 hours and, in 1504, Henry VII ordered the Bankside stews to be temporarily closed.[26] Kohn (ibid.) notes that the Spanish did not give any particular name to the disease but believed it had originated from Native Americans.

The question of an evidence-based link between the disease and a location in the period is best described, therefore, as a question of perspective, not of scientific certainty. Nonetheless, we might hypothesise, if this account of the introduction of the disease is credible, that the disease may be perceived by an English observer as having been associated more with France and Naples than anywhere else. To explore whether the corpus data supports this hypothesis, Figure 2 shows how many times names for venereal disease associated with a particular country appeared in the EEBO-TCP corpus during the seventeenth century as a whole. Note that the chart's scale is logarithmic, and the data is presented in terms of raw rather than normalised frequencies to give readers insight into the actual numbers of mentions of these terms in the corpus. Figure 2 reveals a striking difference between two countries frequently associated with the disease and a handful of other countries infrequently linked to the disease. While "French" venereal disease is mentioned 1,948 times. and "Neapolitan" venereal disease 75 times, no other country is associated with venereal disease in anything other than single figures, with "Scottish" venereal disease being the least common with only two mentions.[27] If these names were inspired by the contemporary belief that the disease was spread into Europe by soldiers of Charles VIII, it appears that English writers did acknowledge the role of Naples as a source of the illness, though they perhaps also regarded the French as being its primary carriers.

---

26. See Burford (1973: 123–24).

27. Both "Scottish" terms take the form of "Scottish fleas" and appear in works by the poet John Taylor (1630 and 1654).

Nationality-driven names for venereal disease in the 17th-century EEBO-TCP corpus
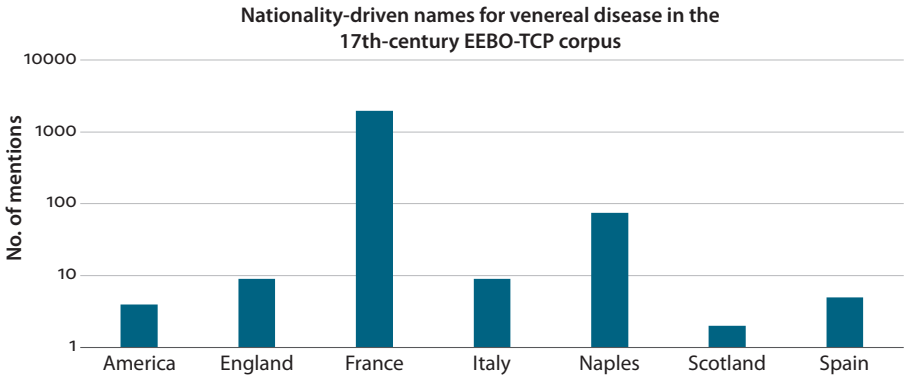
Figure 2. Distribution of seventeenth-century venereal disease nationality-driven labels in the EEBO-TCP corpus. Figures are presented in raw frequencies

Figure 2 is useful in indicating which countries were most often referenced in nationality-driven labels for venereal disease across the entire seventeenth century, but it does not show how frequencies are distributed throughout each decade of the century. Figure 3 isolates the three highest-ranking countries in Figure 2 and shows this distribution per decade.

Numbers of nationality-driven names for venereal disease, France, Naples and Spain
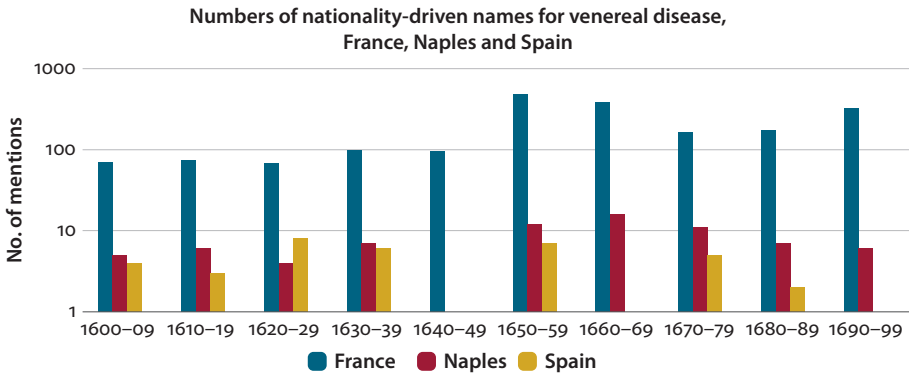
Figure 3. Distribution of seventeenth-century venereal disease nationality-driven names per decade in the EEBO-TCP corpus

If we were to view Figure 3 as a stacked column chart, we would find that the decades in which each of these three terms were most frequently used was the 1650s, 1660s and 1690s. It is possible that the information presented in Figure 3 has been subjected to some level of skew due to the differences in corpus size in each of the decades of the EEBO-TCP corpus, outlined in Table 1. For example, because more words were printed in the 1650s than in any other decade, might this account for the peak in usage of names for venereal disease in that decade? In this case, the

answer is no – when we considered frequency per million words in each decade, the overall picture remained unchanged.

It is tempting to hypothesise why the 1650s, 1660s and 1690s show peaks. Might it be explained by the changing political and social landscape in England, e.g. the period of the Commonwealth in the 1650s with its accompanying shift in social mores or William II's accession to the throne in 1689, which coincided with the ascension of the informal social police force, the Society for the Reformation of Manners? Or might the answer lie in shifting Anglo-French relations? By the middle of the century, France had consolidated its position as the foremost European power and its assertive monarch, Louis XIV, regarded military power as an instrument of foreign policy in order both to defend France's current possessions and to target new acquisitions.[28] Moreover, between 1688 and 1697, England, participating in a wider European coalition, was in direct conflict with France in the Nine Years' War. It is possible that the English population responded to this political threat by increasing its use of demeaning names such as *French pox* or *French disease*.

The work of Ullman (1957: 178, 212), concerning the nature and causes of historical semantic change, may shed further light upon these ideas. Ullman writes that awareness and intention play an important part in change of meaning and acknowledges that such change can also be brought about due to a shift in emotive attitudes towards the referent. Ullman (1962: 232) explains how some loan words have been "filled… with a derogatory sense", giving the example of the term *bougre*, which was used to mean *heretic* and referenced a sect of heretics living in Bulgaria in the Middle Ages, but which then came to mean *sodomite*.[29] In a similar way, it is likely that pejorative connotations stemming from xenophobia have also had an impact upon the dissemination of certain word combinations which start with the term *French*. An etymology dictionary notes that later terms which use the term *French*, such as *French kiss* and *French letter*, both used from the nineteenth century on, probably stemmed from a stereotypical assumption of Gallic sexual sophistication. Other such word combinations, for example *French braid*, also used from the nineteenth century, are less impolite but do still perhaps borrow upon a notion of French worldliness.[30]

Green's *Dictionary of Slang* (2020) goes further, noting the longstanding Anglo-Saxon assumption that French people are "sex-obsessed", particularly when it comes to sexual activity that is perceived to be of an improper nature. Green supports this contention with a seemingly endless list of "French" terms which relate to sexual

---

28. See Tapié (1976: 5).

29. Also see McEnery and Baker (2017b: 207, n.31).

30. See Online Etymology Dictionary https://www.etymonline.com and The Historical Thesaurus of the Oxford English Dictionary http://www.oed.com/thesaurus.

intercourse. However, it is significant that whilst the "French" venereal disease terms date to the sixteenth and seventeenth centuries, those which relate to prostitution and fellatio were used later, from the nineteenth century, and those which relate to male-to-male sex mostly made their appearance in the late twentieth century.[31]

We analysed collocates of the terms *France* and *French* in the EEBO-TCP corpus to find out whether seventeenth-century English writers tended to identify France with general sexual activity. After examining the 500 highest-ranking collocates of both terms, we concluded that France was not particularly associated with sex beyond the topic of venereal disease during this period. Collocates of the term *France* in the seventeenth-century data reveal words which overwhelmingly relate to places, people and foreign policy. The collocates of *French* are more varied, with a few terms relating to food and drink (*brandies, steak, wines, brandy, kidney* [beans], *stew*), language (*idiom, grammar*) and literature (*romance/s, farce, chronicles*). Two collocates in this selection did refer to venereal disease (*pocks, pox*) but there was little else which directly pertained to sex.

In order to explore the ways in which seventeenth-century writers used "French" venereal disease terms, it will be helpful to look more closely at the types of texts in which "French" venereal disease frequently featured. First, however, let us consider the kinds of "French" terms that appear in the EEBO-TCP corpus, as shown in Figure 4.

Figure 4 includes different arrangements of the terms we encountered, including spelling variants. For instance, we have amalgamated the occurrences of *French disease* (537 mentions) with the less common *disease of France* (5 mentions). Likewise, *morbus gallicus* was sometimes expressed as *mobus gallicus, morbus galicus* or *gallicus morbus*. In the case of the term *French goods*, it was necessary to undertake a manual count of occurrences – out of 67 mentions of the term, we decided that only one probably referred to venereal disease (Butler, 1684) and even this was a little ambiguous.

As we can see from Figure 4, the terms *French pox* followed by *French disease* and *morbus gallicus* are the most frequent terms in our study of the EEBO-TCP corpus in the seventeenth century. Figure 5 shows how frequently these three terms were used across the decades in the seventeenth century.

---

31. A number of these terms were taken from two slang dictionaries, Maurer (1981) and Rodgers (1972).

"French" names for venereal disease in the EEBO-TCP corpus, 1600–99

	1	10	100	1000	10000

french cannibal
French curse
french disease
french distemper
french evil
French fever
french goods
french gout
french lues
french malady
french marbles
french pox
french razor
french scab
french sickness
lues gallica
malady of France
morbus gallicus
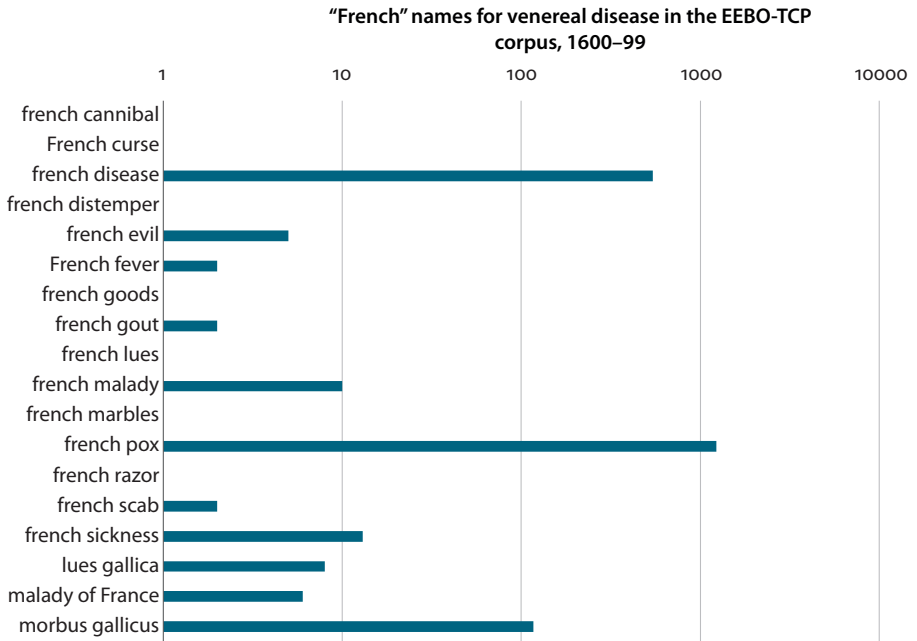
Figure 4. Distribution of names for "French" venereal disease in the seventeenth century in the EEBO-TCP corpus. Figures include spelling variants encountered. Note that those terms which are not represented by bars achieved only one mention in the corpus

Distribution of the three "French" most popular venereal disease names across the 17th century

No. of mentions

	1600–09	1610–19	1620–29	1630–39	1630–39	1650–59	1660–69	1670–79	1680–89	1690–99

● french disease   ● french pox   ● morbus gallicus

Figure 5. Usage of *French disease, French pox* and *morbus gallicus* throughout the decades of the seventeenth century in the EEBO-TCP corpus. Variant forms and spellings were included

These three terms follow a not dissimilar pattern to the one presented in Figure 3, peaking in the middle of the century and then again in the 1690s. Indeed, when we consider all the "French" terms combined, it is clear that they burgeoned in popularity in the 1650s (2.84 mentions per million words in that decade) and continued to be used with very high frequencies in the 1660s (3.46 menti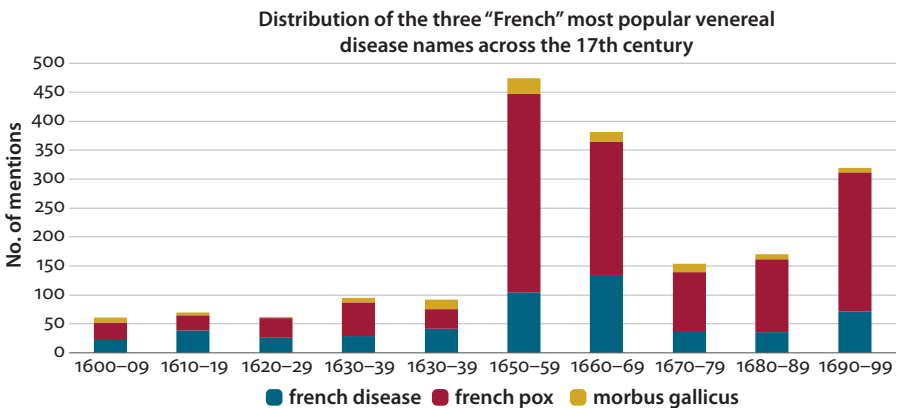ons per million words). A drop in frequency in the 1670s and 1680s was followed by another peak in the 1690s (2.51 mentions per million words).

Figure 5 reveals that *morbus gallicus* (and its variants) was used far less frequently and its usage was more stable. Qualtiere and Slights (2003: 5) were correct in their belief that *French pox* was commonly used but their argument that *morbus gallicus* was most frequently used towards the middle of the sixteenth century is less convincing.[32] The other names present in Figure 4, such as *disease of France* (5 hits over the century), *French fever* (2 hits), *French gout* (2 hits), *French distemper* (1 hit) and *French lues* (1 hit), feature in EEBO-TCP but with very low frequencies. The occurrence of *French sickness* is slightly higher with 13 hits but 7 of these are attributed to one medical text from the 1670s (Croll 1670). The terms *French bone-ache, French chilblains, French crust, French measles* and *French plague* which were encountered during our initial lexicographical research and background reading do not appear in the EEBO-TCP corpus.

When authors employed non-"French" nationality terms, e.g. *Neapolitan disease* or *American disease*, they often went on to specifically discuss the etymology or the origin of the disease. Such discussions took place in 60 per cent of instances when Naples was linked to venereal disease.[33] However, when writers referenced "French" venereal disease, they did not feel the need to justify why they were attributing the disease to France, e.g. "he says it cures the French Pox, Leprosy, Gout, and Falling sickness" (Primerose, 1640). In other words, the names *French pox* or *French disease* were an accepted part of the lexicon and did not require explanation.

3.2   Genre distribution

So far, we have shown that, in the "geography of names", venereal disease was far more likely to be associated with France than any other country. The next step is to explore the kinds of works in which "French" venereal disease featured in order to throw light upon how it was used by seventeenth-century writers.

---

32. *Morbus gallicus* achieves a frequency of 0.13 instances per million words in the sixteenth-century part of the EEBO-TCP corpus.

33. See, for instance, Gayton (1659).

Table 2 shows the distribution by genre for all matches of "French" venereal disease terminology that feature in Figure 4. Raw frequencies of matches are listed alongside frequency per million words in parentheses – the latter figures were calculated relative to the total numbers of words available for each decade of the corpus. By focussing on frequency per million words calculated in this way, we could ensure that our results were not influenced by changes in corpus size throughout the decades. Table 2 only lists the six genres in which "French" venereal disease terms achieved a frequency of over 0.1 per million in at least one decade of the seventeenth century. This is, admittedly, an arbitrary rule of thumb but we felt that it permitted us to look at cases which offered sufficient evidence.[34]

Table 2. Genre distribution for "French" venereal disease names

	1600s	1610s	1620s	1630s	1640s	1650s	1660s	1670s	1680s	1690s
Astrology and Predictions	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	36 (0.21)	0 (0)	5 (0.04)	6 (0.01)	0 (0)
Declarations	0 (0)	5 (0.08)	8 (0.14)	0 (0)	0 (0)	2 (0.01)	0 (0)	1 (0.01)	0 (0)	0 (0)
History	5 (0.09)	7 (0.11)	8 (0.14)	7 (0.11)	1 (0.01)	15 (0.09)	7 (0.06)	10 (0.08)	9 (0.06)	7 (0.05)
Philosophy	2 (0.03)	0 (0)	21 (0.38)	2 (0.03)	2 (0.02)	3 (0.02)	3 (0.03)	4 (0.03)	1 (0.01)	5 (0.04)
Science: Medical	10 (0.17)	25 (0.4)	9 (0.16)	53 (0.83)	39 (0.45)	147 (0.87)	110 (0.98)	46 (0.39)	73 (0.51)	213 (1.66)
Treatise	28 (0.49)	7 (0.11)	0 (0)	9 (0.14)	16 (0.18)	20 (0.12)	76 (0.68)	12 (0.1)	25 (0.18)	47 (0.37)

In order to discover if the peaks in each row simply corresponded with periods when more texts categorised as particular genres were printed, it is helpful to look at how the size of each of these genres varies across the century. Table 3 shows the size of the six genres identified in Table 2. It is clear that there is little relation between the size of genre and how many times texts within each genre mention "French" venereal disease. For example, the genre of History is at its largest in the 1650s, 1680s and 1690s, yet references to "French" venereal disease in texts categorised as this genre peak in the 1620s. Likewise, there is a massive increase in texts classified as Science: Medical which reference the illness in the 1690s, but the genre is at its largest in the 1630s.

---

34. Note that the Unclassified category was not considered in this study.

Table 3. Genre size in the EEBO-TCP corpus across the seventeenth century

Decade	Astrology and predictions	Declarations	History	Philosophy	Science: Medical	Treatise
1600–09	258,684	1,720,852	6,872,066	564,922	2,138,334	8,910,989
1610–19	143,571	1,260,459	8,817,541	1,240,219	2,352,999	3,946,489
1620–29	138,982	838,486	5,892,172	3,739,012	1,074,044	4,696,246
1630–39	13,376	780,653	5,276,899	1,964,690	4,811,867	6,184,596
1640–49	71,672	1,541,837	4,863,223	769,192	1,634,410	4,982,705
1650–59	1,296,480	1,229,874	16,286,277	2,150,047	3,350,178	8,575,152
1660–69	681,957	812,787	10,969,330	2,617,194	3,828,883	7,266,617
1670–79	468,743	1,438,149	13,249,008	1,378,235	2,681,931	6,115,644
1680–89	890,482	1,274,625	16,178,882	1,936,611	4,075,936	8,881,706
1690–99	230,457	526,582	16,659,216	2,707,006	3,323,605	6,923,390

Let us explore the genres which show peaks in the early decades of the century. Declarations (peaks in the 1610s and 1620s); Philosophy (peaks in the 1620s); and History (peaks in the period 1610–39) do have larger frequencies in particular decades, but these are relatively moderate. Texts categorised as belonging to the History genre discuss the origin of venereal disease and its impact on historical nations (e.g., Fage (1667) and Purchas (1613)). The peaks in the 1610s and 1620s for Declarations are attributable to only two texts, J.T. (1619) and Fioravanti (1626). Similarly, the peak in the genre of Philosophy in the 1620s can be explained by repeated references to *French pox* and the *French disease* by Samuel Purchas in his popular collection of travel stories, *Purchas his Pilgrimes* (1625). Purchas discussed how venereal disease was brought into Africa and how Egypt was the worst affected country. He also described the kinds of medicines that were used in places such as Pegu in present-day Myanmar:

> …the greatest merchandise there is Verzin and Nypa, which is an excellent Wine, which is made of the flower of a tree called Nyper. Whose liquor they distill, and so make an excellent drink clear as Crystal, good to the mouth, and better to the stomach, and it hath an excellent gentle virtue, that if one were rotten with the French pox, drinking good store of this, he shall be whole again.
>
> (Purchas, *Purchas his Pilgrimes*, 1625)

Taavitsainen (2001a: 142) has emphasised the importance of understanding the purpose of the text when considering genre. Although it appears that Purchas had no ulterior motives other than in educating his readers about far-off cultures they would almost certainly never experience first-hand, his publication of 1625 contained a great deal of material from Richard Hakluyt. Hakluyt's writing was intended to promote interest in the colonisation of North America. Indeed, Bailey-Goldschmidt and Kalfatovic (1993: 142) have argued that from the seventeenth century, travel

accounts were increasingly "tied up with the larger production of knowledge from the period, and the rise of European hegemony".

The genre Astrology and Predictions is unique because it peaks in the middle of the century only, in the 1650s. The relevant texts discuss the origins of venereal disease in nations and individuals based on the alignment of planets and treatments for the illness. For instance, Ramesey (1653) explained:

> …an Eclipse of the Sun in the first Face of Cancer, troubles the air, and causes varous winds and alterations of weather. In the second, dries up Rivers and Fountains, and causes petulant and grievous mortalities. In the third, through Armenia and Africa, and the rest of the Regions and Places subject to Cancer, Hidropical diseases, seditions, and the disease of France, Naples, Spain, England, and indeed of all the World; yet commonly called the French and Neapolitan disease, or more plainly the French Pox. (Ramesey, *Astrologia Restaurata*, 1653)

Nicholas Culpeper, one of the most well-known medical astrologers of the early modern period, published texts in 1651 and 1652 in which he repeatedly referenced "French" venereal disease and how it might be treated. Interestingly, Culpeper (1651) included French pox in a categorisation of illnesses attributed to Venus but did not ascribe it to "inordinate love or lust", as he did with the conditions of priapism and impotency.

The genre of Treatise also takes an unusual pathway as mentions of venereal disease in its texts peak in the decades 1600–09 and 1660–69 and, less strikingly, 1680–99. In the first decade of the century, texts about "French" venereal disease belonging to the Treatise genre span a range of subjects, such as anti-Catholic rhetoric (e.g. Valera, 1600) and geographical histories of far-off countries in Africa and Asia, e.g. "That wood which is of the brown or tawny colour is used by the African physicians for the curing of the French pox, whereupon it is commonly called by the name of pock-wood." (Leo, 1600).

The wide range of topics discussed with reference to venereal disease continued into the 1660s and 1680s. The German physician, Daniel Sennert (1660) introduced his substantial treatise dedicated to *venereal pocks* by discussing the origin of the disease. He asserted that the disease appeared in French camps during warfare between France and Naples in the late fifteenth century, "and therefore by the Italians was first of all named the French disease". Browne (1684) did not entirely agree: "The French Pox no new Disease. which although it was formerly warmly treated at Naples, and had a plentiful entertainment in the West-Indies, yet it had being in the World long before this, as History testifies".

A volume by the legal writer, William Sheppard (1662) takes us to the topic of litigation. He explained that accusing a man of having a "dangerous disease, by reason whereof he ought to separate himself, or be separated by the Law from

the society of men", such as French pox or plague, is actionable. By juxtaposing plague with pox and suggesting that sufferers of both diseases might require social isolation, Sheppard gives us an insight into how venereal disease continued to be perceived into the middle of the century.

Accusations and insults about venereal disease are also highlighted in the Trials and Disputes genre, which did not achieve a frequency of 0.1 in at least one decade of the century but did display a very modest peak in the 1650s. Accusing a real person of having "French" venereal disease was a serious business and could lead to the loss of reputation for the accused and official sanctions for the accuser. Verdicts showed awareness of whether or not the term *pox* was used to specifically indicate venereal disease and the term *French pox* was deemed to be a useful indictor of this. For example, Glisson et al. (1679) concluded that the accusation "M is a Whore and hath the Pox, Pox. and they have made such holes in her that you may turn your finger in them, and K the Surgeon hath given her a Diet drink to heal them, take heed how you drink with her" alluded to French pox rather than small pox because i.) whores cause pox; ii.) a treatment for French pox was mentioned and iii.) a warning about contagion suggests the French pox. This corresponds to earlier research (McEnery and Baker, 2016) which uncovered an early modern association between venereal disease and prostitution and a determination among women in particular to pursue defamation lawsuits when they felt their reputation had been compromised.[35] Interestingly, the use of "French" venereal disease as an insult also popped up in another low frequency genre – Play: Comedy. However, the small numbers of relevant mentions in that genre merely refer to the illness as one of many that quack doctors claim to cure.[36]

Returning to the genre of Treatise, Johnson (1665) goes one step further in explaining why a sufferer might attempt to conceal their symptoms: "The French Disease, having so much of shame intayled upon it, because Contracted usually from great Debaucheries, it is the great care of all Persons punished with it, to conceal it as much as is possible". Although Fraser (1984: 411) has asserted that the illness was so prevalent by the end of the fifteenth century that it achieved a certain amount of social tolerance, the corpus suggests that symptoms of venereal disease were considered best concealed well into the seventeenth century. This accords with the view of Siena (1998: 272), who has argued that the illness was accompanied by a considerable level of social stigma.[37] Johnson (ibid.) goes on to

---

35. See also Sharpe (1980: 15,17) and Gowing (1993: 2–4, 15).

36. See, for example, Thomson (1668).

37. Sumich (2013: 142) has drawn attention to entries in Samuel Pepys' diaries which related to his concerns over the possibility of his brother having contracted venereal disease and his wife's desire to conceal her treatment for a vaginal infection in case this resulted in gossip about the nature of her illness.

comment that the disease which most quack medical practitioners falsely claim to have cured to the greatest extent is that of Lues Venera.[38] Harrington (1682) agrees the "noisome French-Disease" comes "of filthy Lust". Qualtiere and Slights (2003) have argued that perceptions of French pox shifted between the sixteenth and seventeenth centuries, from the illness being attributed to God's wrath to being blamed upon individual moral failings.

Now let us turn our attention to the genre which displays a massive peak at the very end of the seventeenth century, Science: Medical. Perhaps unsurprisingly, these texts deal with the symptoms and treatments of venereal disease. Bate et al. (1694) is a common example of these texts, being a translation of an earlier text by the physician George Bate, who treated Charles II. It suggests a number of treatments for French pox, including "The Magistery of Turbith Mineral", which appears to have been sulfate of mercury,[39] Guajacum,[40] and "the acid Spirit of Vitriol" (sulphuric acid). The volume also described some of the symptoms of the disease, such as "Crystaline Bladder" (bladder stones), "Pocky Pustles and Sores", "pocky Nodes" and "Veneral Buboes".

Although there are 213 references to "French" venereal disease in the Science: Medical genre in the 1690s, these references only appear in eight texts. Moreover, four of these texts (Pechey 1692, 1694a, 1694b and 1697) were authored by John Pechey, a licensed practitioner who advertised his services from a private practice in London.[41] Pechey offered a service whereby patients could send a completed questionnaire pertaining to their symptoms to him, and he would send a diagnosis and appropriate medicine in the return post. His advertisements for pills to cure scurvy and French pox provoked the ire of the Royal College of Physicians.[42] However, these learned men may not have been entirely motivated by a concern for patients – Furdell (2002: 149) notes that trained physicians disapproved of sharing "secrets of the trade". Although these texts do not focus exclusively on venereal disease but instead described treatments suitable for a wide range of illnesses, references to venereal disease were scattered throughout them repeatedly. For instance, Pechey's (1694a) helpful index entry for *French pox* directed readers to ten separate page

---

38. Siena (2001) has written about the prevalence of non-licensed medical practitioners in London who offered discreet treatments for venereal disease.

39. See Eastaugh et al. (2008).

40. Gum guiacum had been used as a treatment for venereal disease since the sixteenth century. The late sixteenth-century Flemish print *Hyacum et lues venera* depicts the servants of a convalescent man preparing the medicine. Also see Churchill (2005: 8) for a discussion of contemporary treatments.

41. Pechey (1692) is mostly taken from the writings of Thomas Sydenham.

42. See Pechey, John, Oxford DNB: https://doi.org/10.1093/ref:odnb/21737.

numbers. The entry for French pox was preceded by one for Freckles and followed by one for Frenzy, but each of these is only allotted a single page number.

These medical texts do not focus on the origins of the disease or how it is contracted. Rather, the authors usually maintained an objective gaze, neither blaming nor shunning the sufferer, and their response to the disease was intended to be constructive. However, we cannot go so far as to suggest that these volumes took a compassionate approach to healing venereal disease. In almost every text which suggested treatments, there was one important omission – that of the patient themselves. The texts talked of "the cure of the French-Pox" (Bate et al., 1694) and how certain medicines were "good for the French-Pox" (Pechey, 1694a) or "given in the French Pox" (Salmon, 1698), but there was very little reference to the experiences of the sufferer and no attempt to tailor treatment for specific groups of patients. Perhaps most reckless was the lack of transparency about the side effects that some of these suggested treatments would entail.

4.   Conclusion

This study has explored how terminology referring to venereal disease was constructed in seventeenth-century texts in the EEBO-TCP corpus. In the first section we detailed the methods we used in order to arrive at a useful collection of query terms. We decided not to base our study on the term *pox* because, in a substantial proportion of cases, it did not refer to venereal disease. However, the vast majority of words which we believed may have been used to refer to venereal disease by seventeenth-century writers were, in fact, not present in the EEBO-TCP corpus or were at least used very rarely. When faced with the conundrum of either studying a very high-frequency polysemous term (*pox*) or analysing a moderately frequent term (primarily *French pox*), we decided on the latter and also used the opportunity to include an exploration of other terms for the disease associated with foreign countries. We found that English writers were far more likely to attribute venereal disease to their immediate neighbour, France. Naples attained second ranking, but references to "Neapolitan" venereal disease were much lower than references to "French" venereal disease.

Our genre analysis showed that "French" venereal disease tended to be mentioned more frequently in a fairly small number of genres. It was often referenced in discussion to other cultures, particularly in the first half of the century, not least in debates regarding whether or not the disease originated from a particular nation. The illness was, at least, clearly a cause for embarrassment for its sufferers. At worst, it led to social ostracism for them. Such was the potency of an accusation that somebody had the illness that defamation cases involving insults concerning

venereal disease are alluded to in the corpus. It is also clear that sufferers of the disease were regarded by some as having brought it upon themselves through unclean living. Such judgements were more apparent in the earlier works of the corpus but continued right up until the end of the seventeenth century. We also found an association between prostitution and venereal disease, with the collocates of POX leading us to the nickname, *the whore's disease*. However, the claim in previous research literature of a wider belief that venereal disease was primarily spread by women, written about by Siena (1998), Spongberg (1997) and Spates (2006), was less apparent.

The most noticeable shift in genre distribution occurred at the very end of the century, with a sudden upsurge in references to venereal disease in medical texts. Here, readers were provided with some stark details of symptoms and a plethora of worrying treatments which writers confidently asserted would cure the disease. These texts offered a new tone to the narrative surrounding venereal disease. Judgement and blame are usually absent from these texts, indeed they tend to omit any references to the human sufferer altogether; instead they focus on constructive methods to treat or cure the illness which, in a period where the disease brought great shame, may have been welcome to those experiencing distressing symptoms.

Acknowledgements

References

Primary sources[43]

B.E. 1699. *A New Dictionary of the Canting Crew in its Several Tribes of Gypsies, Beggers, Thieves, Cheats &c*. London.

Bate, George, James Shipton and William Salmon. 1694. *Pharmacopœia Bateana, or, Bate's dispensatory*. London.

Browne, John. 1684. *Adenochoiradelogia*. London.

Butler, Samuel. 1684. *Hudibras in three parts*. London.

Cavendish, William. 1649. *The Country Captained and the Varietie*. London.

Croll, Oswald. 1670. *Bazilica Chymica, & Praxis Chymiatricæ*. London.

Culpeper, Nicholas. 1651. *Semeiotica Uranica, or, An Astrological Judgment of Diseases from the Decumbiture of the Sick*. London.

---

43. Titles have been shortened due to space constraints.

Culpeper, Nicholas. 1652. *The English Physitian*. London.

Fage, Robert. 1667. *Cosmography or, a Description of the Whole World*. London.

Fioravanti, Leonardo. 1626. *A Discourse vpon Chyrurgery*. London.

Gayton, Edmund. 1659. [no title]. London.

Glisson, William, Anthony Gulston, William Style and Henry Applegarth. 1679. *The Common Law Epitomiz'd with Directions How to Prosecute and Defend Personal Actions, Very Useful for All Lawyers, Justices of Peace, and Gentlemen*. London.

Harrington, James. 1682. *Horæ Consecratæ*. London.

J.T. 1619. *The Hunting of the Pox a Pleasant Discourse betweene the Authour, and Pild-Garlicke*. London.

Johnson, William. 1665. *Agyrto-Mastix*. London.

Leo, Africanus. 1600. *A Geographical Historie of Africa*. London.

Lowe, Peter. 1569. *Easie, certaintie, and perfect method, to cure and preuent the Spanish sicknes*. London.

Marston, John. 1598. *The Metamorphosis of Pigmalions Image*. London.

Pechey, John. 1692. *A Collection of Chronical Diseases viz. the Colick, the Bilious Colick, Hysterick Diseases, the Gout, and the Bloody Urine from the Stone in the Kidnies*. London.

Pechey, John. 1694a. *The Compleat Herbal of Physical Plants Containing All Such English and Foreign Herbs, Shrubs and Trees as are Used in Physick and Surgery*. London.

Pechey, John. 1694b. *The London Dispensatory, Reduced to the Practice of the London Physicians wherein are Contain'd the Medicines, Both Galenical and Chymical, that are Now in Use*. London.

Pechey, John. 1697. *A Plain Introduction to the Art of Physick Containing the Fundamentals, and Necessary Preliminaries to Practice*. London.

Primerose, James. 1640. *The Antimoniall Cup Twice Cast*. London.

Purchas, Samuel. 1613. *Purchas his Pilgrimage*. London.

Purchas, Samuel. 1625. *Purchas his Pilgrimes In Fiue Bookes*. London.

Ramesey, William. 1653. *Astrologia Restaurata*. London.

Salmon, William. 1698. *Ars Chirurgica a Compendium of the Theory and Practice of Chirurgery in Seven Books*. London.

Sennert, Daniel. 1660. *Two Treatises*. London.

Sheppard, William. 1662. *Action upon the Case for Slander*. London.

Taylor, John. 1630. *All the vvorkes of Iohn Taylor the water poet*. London.

Taylor, John. 1654. [untitled]. London.

Thomson, Thomas. 1668. *The English Rogue*. London.

Valera, Cipriano de. 1600. *Two Treatises*. London.

Secondary sources

Bailey-Goldschmidt, Janice, and Martin Kalfatovic. 1993. "Sex, Lies and European Hegemony: Travel Literature and Ideology." *Journal of Popular Culture* 26 (4): 141–153. https://doi.org/10.1111/j.0022-3840.1993.2604_141.x

Bergs, Alexander. 2004. "Letters: A New Approach to Text Typology." *Journal of Historical Pragmatics* 5 (2): 207–227. https://doi.org/10.1075/jhp.5.2.04ber

Burford, Ephraim John. 1973. *The Orrible Synne: A Look at London Lechery from Roman to Cromwellian Times*. London: Calder and Boyars.

Churchill, Wendy. 2005. "The Medical Practice of the Sexed Body: Women, Men, and Disease in Britain, circa 1600–1740." *Social History of Medicine* 18 (1): 3–22. https://doi.org/10.1093/sochis/hki006

Cohen, Ralph. 1986. "History and Genre." *Neohelicon* 13 (2): 87–105. https://doi.org/10.1007/BF02028902

Eastaugh, Nicholas, Valentine Walsh, Tracey Chaplin and Ruth Siddall. 2008. *Pigment Compendium: A Dictionary and Optical Microscopy of Historical Pigments*. London and New York: Routledge. https://doi.org/10.4324/9780080943596

Fraser, Antonia. 1984. *The Weaker Vessel: Woman's Lot in Seventeenth Century England*. London: Weidenfeld & Nicolson.

Furdell, Elizabeth. 2002. *Publishing and Medicine in Early Modern England*. New York: University of Rochester Press.

Gowing, Laura. 1993. "Gender and the Language of Insult in Early Modern London." *History Workshop Journal* 35: 1–21. https://doi.org/10.1093/hwj/35.1.1

Green, Jonathan. 2020. *Green's Dictionary of Slang*. Digital Edition. https://greensdictofslang.com/

Hamlin, William. 2014. "God-language and Scepticism in Early Modern England: An Exploratory Study Using Corpus Linguistics Analysis as a Form of Distant Reading." *English Literature* 1 (1): 17–41.

Hardie, Andrew. 2012. "CQPweb – Combining Power, Flexibility and Usability in a Corpus Analysis Tool." *International Journal of Corpus Linguistics* 17 (3): 380–409. https://doi.org/10.1075/ijcl.17.3.04har

Kohn, George (ed.) 2008. *Encyclopedia of Plague and Pestilence: From Ancient Times to the Present*. New York: Facts on File.

Lancashire, Ian. 2015. *Lexicons of Early Modern English*. Toronto: University of Toronto Library and University of Toronto Press. https://leme.library.utoronto.ca/

Linane, Fergus. 2003. *London: The Wicked City*. London: Robson.

Maurer, David. 1981. *Language of the Underworld*. Lexington, Kentucky.

McEnery, Tony, and Helen Baker. 2016. *Corpus Linguistics and 17th-Century Prostitution: Computational Linguistics and History*. Bloomsbury: London.

McEnery, Tony, and Helen Baker. 2017a. "Language Surrounding Poverty in Early Modern England: A Corpus-Based Investigation of How People Living in the Seventeenth-Century Perceived the Criminalised Poor." In *From Data to Evidence in English Language Research*, ed. by Carla Suhr, Terttu Nevalainen and Irma Taavitsainen, 225–257. Leiden and Boston: Brill.

McEnery, Tony, and Helen Baker. 2017b. "The Public Representation of Homosexual Men in Seventeenth-Century England: A Corpus Based View." *Journal of Historical Sociolinguistics* 3 (2): 197–217. https://doi.org/10.1515/jhsl-2017-1003

Murphy, Sean. 2019. "Shakespeare and His Contemporaries: Designing a Genre Classification scheme for Early English Books Online 1560–1640." *ICAME Journal* 43: 59–82. https://doi.org/10.2478/icame-2019-0003

Orr, Leah. 2011. "Genre Labels on the Title Pages of English Fiction, 1660–1800." *Philological Quarterly* 90 (1): 65–95.

Qualtiere, Louis F. and William W. E. Slights. 2003. "Contagion and Blame in Early Modern England: The Case of the French Pox." *Literature and Medicine* 22 (1): 1–24. https://doi.org/10.1353/lm.2003.0009

Pechey, John. *Oxford Dictionary of National Biography* (online ed.). Oxford University Press. https://doi.org/10.1093/ref:odnb/21737

Ratia, Maura. 2013. "Investigating Genre through Title-Pages: Plague Treatises of the Stuart Period in Focus." In *Principles and Practices for the Digital Editing and Annotation of Diachronic Data*, ed. by Anneli Meurman-Solin and Jukka Tyrkkö. Helsinki: Varieng. https://varieng.helsinki.fi/series/volumes/14/ratia/

Rodgers, Bruce. 1972. *The Queens' Vernacular*. San Francisco, California: Straight Arrow Books.

Sharpe, James A. 1980. *Defamation and Sexual Slander in Early Modern England: The Church Courts at York*. York: Borthwick. https://doi.org/10.3138/9781442679856-002

Siena, Kevin. 1998. "Pollution, Promiscuity, and the Pox: English Venereology and the Early Modern Medical Discourse on Social and Sexual Danger." *Journal of the History of Sexuality* 8 (4): 553–574.

Siena, Kevin. 2001. "The 'Foul Disease' and Privacy: The Effects of Venereal Disease and Patient Demand on the Medical Marketplace in Early Modern London." *Bulletin of the History of Medicine* 75 (2): 199–224. https://doi.org/10.1353/bhm.2001.0097

Smitterberg, Erik, and Merja Kytö. 2015. "English Genres in Diachronic Corpus Linguistics." In *From Clerks to Corpora: Essays on the English Language Yesterday and Today*, ed. by Philip Shaw, Britt Erman, Gunnel Melchers and Peter Sundkvist, 117–133. Stockholm: Stockholm University Press. https://doi.org/10.16993/bab.g

Spates, William. 2006. "Proverbs, Pox, and the Early Modern Femme Fatale." *Notes and Queries* 53 (1): 47. https://doi.org/10.1093/notesj/gjj124

Spongberg, Mary. 1997. *Feminizing Venereal Disease: The Body of the Prostitute in Nineteenth-Century Medical Discourse*. New York: New York University Press. https://doi.org/10.1057/9780230375130

Sullivan, Ceri. 2007. "Disposable Elements? Indications of Genre in Early Modern Titles." *The Modern Language Review* 32 (2): 641–653. https://doi.org/10.2307/20467425

Sumich, Christi. 2013. *Divine Doctors and Dreadful Distempers: How Practicing Medicine Became a Respectable Profession*. Amsterdam and New York: Rodopi. https://doi.org/10.1163/9789401209472

Szreter, Simon. 2017. "Treatment Rates for the Pox in Early Modern England: A Comparative Estimate of the Prevalence of Syphilis in the City of Chester and its Rural Vicinity in the 1770s." *Continuity and Change* 32 (2): 183–223. https://doi.org/10.1017/S0268416017000212

Taavitsainen, Irma. 2001a. "Changing Conventions of Writing: The Dynamics of Genres, Text Types, and Text Traditions." *European Journal of English Studies* 5 (2): 139–150. https://doi.org/10.1076/ejes.5.2.139.7309

Taavitsainen, Irma. 2001b. "Middle English recipes: Genre characteristics, Text Type Features and Underlying Traditions of Writing." *Journal of Historical Pragmatics* 2 (1): 85–113. https://doi.org/10.1075/jhp.2.1.05taa

Tapié, Victor. 1976. "Louis XIV's Methods in Foreign Policy." In *Louis XIV and Europe*, ed. by Ragnhild Marie Hatton, 3–15. London: Macmillan. https://doi.org/10.1007/978-1-349-15659-7_1

Ullman, Stephen. 1957. *The Principles of Semantics*. Oxford: Basil Blackwell.

Ullman, Stephen. 1962. *Semantics: An Introduction to the Science of Meaning*. Oxford: Basil Blackwell.

Vazquez, Maegan. 2020. "Trump says he's pulling back from calling novel coronavirus the 'China virus'". CNN. 24 March 2020. https://edition.cnn.com/2020/03/24/politics/donald-trump-pull-back-coronavirus-chinese-virus/index.html

Vazquez, Maegan, and Betsy Klein. 2020. "Trump again defends use of the term 'China virus'. *CNN*, 17 March 2020. https://edition.cnn.com/2020/03/17/politics/trump-china-coronavirus/index.html

CHAPTER 3

Medical topics and style from 1500 to 2018
A corpus-driven exploration

Gerold Schneider
University of Zurich

This chapter investigates changes in medical topics, style and language across 500 years, from 1500 to 2018. To do so, we employ data-driven methods of Computational Linguistics and Digital Humanities: document classification, topic modelling, and automatically constructed conceptual maps. We trace changes from traditional thinking in the scholastic period to empirical methods, professionalised medicine, and finally the increasing importance of data, statistics and clinical studies, away from symptom-centred medicine. We conclude that medical discourse has undergone radical changes and that data-driven methods reflect these changes and offer an unprecedented overview. We also critically discuss shortcomings of our data and methods.

Keywords: data-driven approaches, machine learning, collocations, Topic Modelling, history of medicine, Digital Humanities, conceptual maps, Kernel Density Estimation, automated content analysis, English medical discourse, language and health, culturomics

1.   Introduction

The enormous body of scientific and medical writing is "one of the more remarkable of human literary accomplishments" (Bazerman 1988: 13) and one of the outstanding achievements of civilisation and society. Using state-of-the-art quantitative methods, this chapter offers an overview of the patterns of historical development in medical discourse across a 500-year period. Approaches from culturomics and computational linguistics have been applied to automatically extract the most salient information from large English corpora, which is then qualitatively interpreted and assessed.

2.   Motivation

Data science and culturomics employ quantitative methods to uncover trends, based on the assumption that correlations between frequency, topicality and mental salience are very strong. This assumption has repeatedly been confirmed (e.g. Janda 2013, Conklin and Schmitt 2012: 56).

The correlation between topic salience and frequency has also been confirmed in media analysis (e.g. Roberts et al. 2002). Based on agenda-setting studies, Ghanem (1997: 12) argues that "[t]he frequency with which a topic is mentioned probably has a more powerful influence than any particular framing mechanism".

The present study uses culturomics (Michel et al. 2011) as its starting point. The culturomics approach goes back to the compilation of the Google n-gram corpus and its publicly accessible web interface.[1] Its compilers collected over 5 million books, 4% of all printed books, to achieve a hitherto unparalleled coverage. They argue that their approach of mainly analysing peaks of lexical items opens up a new field of research including

> [...] lexicography, the evolution of grammar, collective memory, the adoption of technology, the pursuit of fame, censorship, and historical epidemiology. Culturomics extends the boundaries of rigorous quantitative inquiry to a wide array of new phenomena spanning the social sciences and the humanities.
>
> (Michel et al. 2011: 1)

We aim to go beyond the original culturomics approach in several ways. First, while Michel et at. (2011) reliably find peaks in the data, the lexical items are selected by the researcher, which is demanding and requires theoretical expertise. Second, we also apply a wide range of computational methods. Third, we use carefully manually compiled medical corpora as far as possible. We shall take up these points one by one.

2.1   Systematic comparison of all lexical features

Approaches where all patterns are compared to each other, driven by observable differences in frequencies of patterns, are recognised as a desideratum by Sinclair and Carter:

> We should strive to be open to the patterns observable in language [...] and rebuild a picture of language and meaning which is not only consistent with the evidence but also exploits it to the full [...] the first stage should be an attempt to inspect the data with as little attention as possible to theory.   (Sinclair and Carter 2004: 10)

---

1.   https://books.google.com/ngrams

The term "corpus-driven" (or "data-driven") is attributed to Tognini-Bonelli (2001: 84–91) in corpus linguistics. For her, the advantages of corpus-driven methods include relative neutrality to theory: "It is accepted that there is no such thing as a theory-neutral stance, but in CDL [corpus-driven linguistics] the attempt is made to suppress all received theories, axioms and precepts" (ibid: 178). Corpus-driven approaches thus build up hypotheses in a bottom-up fashion because they pick up any strong difference or correlation, and have the potential to discover new insights (e.g. Hilpert and Gries 2016: 44–45).

Many researchers argue that any type of statistically significant difference, connection or association between not clearly connected entities and facts could be meaningful (see e.g. Ananiadou et al. 2006: 572). In the medical domain, data-driven approaches increasingly complement hypothesis-driven approaches:

> Systems biology is one of the key examples of a field where the mode of scientific knowledge discovery is shifting from a hypothesis-driven mindset to an integrated holistic mode that combines hypotheses with data […]. In the data-rich but hypothesis-poor sciences, including functional genomics and most of biomedicine, the normative hypothesis-driven, deductive scientific method becomes increasingly difficult to sustain because it is unable to deliver advances in knowledge quickly enough.   (Ananiadou et al. 2006: 571)

In fields like history, too, the identification of peaks and trends may be useful e.g. in the detection of the spread of epidemics on social media (Steinberger et al. 2015), and the argument for a holistic view is crucial.

Our wish to systematically detect differences at the lexical level leads to our first RQ.

RQ1.   Can a systematic comparison of quantitative frequencies of terms and linguistic features detect historical developments in the medical domain? Are the reported differences interpretable?

To address this RQ, we use document classification. Our hypothesis is that the strongest features of each century will show us salient terms.

2.2   Advanced computational methods

The present study also applies a wide range of computational methods: along with document classification (see RQ1), we use topic modelling to abstracting from words to concepts, which leads us to our second RQ.

RQ2.   Can the underlying topics and concepts be detected by quantitative methods?

We apply topic modelling to address this RQ. Our hypothesis is that we can detect how topics change over time.

While the above methods have been used in automated media content analysis (Grimmer and Stewart 2013), we also test more experimental and rarely used methods for drawing conceptual maps automatically. RQ3 aims to find out how successful they are.

RQ3.   Can the underlying assumptions, trends in the history of thought, be uncovered?

We automatically draw conceptual maps to explore a hopefully semantically meaningful landscape. From the perspective of the history of thought, we would expect to see a philosophical move from religion to scholastic rigour and empirical research, paired with increased specialisation.

2.3   Sampling and representativeness

Data-driven approaches are known to be susceptible to sampling issues (Tognini-Bonelli 2001: 88). Michel et al. (2011) address this issue by using the largest sample possible, *Google Books*, which covers 4% of all books. The statistically well-founded hope is that the larger the sampling frame is, the smaller the risk becomes that it can be unrepresentative, but Michel et al. (2011) do not report results for specific genres and registers separately.

We aim to go beyond Michel et al. (2011) using carefully manually compiled medical corpora wherever possible. However, as we had to rely on available material, our selection is not ideal, which brings us to RQ4.

RQ4.   How far does our selection of material affect our results? Can we assess its potential impact?

We address how well data-driven approaches work, and what patterns can be discovered in medical discourse over a very long time span from a remote bird's-eye perspective in an exploratory fashion. This approach offers a new perspective on the history of medical discourse, and visualisation techniques show us how radically this domain of writing has changed across the centuries.

3. Materials

Taken together, the corpora used in this study cover the period from 1500 to 2018, and include medical documents from carefully sampled English corpora. They may not be ideally representative as the compilation policies vary and genre drift may cause a skew. However, this can also be seen as an inherent characteristic to measure. The corpora are presented in temporal order.

3.1 CEEM

The three components of *Corpus of Early English Medical Writing* (CEEM)[2] – MEMT, EMEMT and LMEMT – cover the history of English vernacular medical writing up to 1800, from the earliest manuscripts to the beginnings of modern clinical medicine. These corpora have been compiled following the same principles, reflecting the underlying traditions of writing, which enables a diachronic analysis of the underlying genres.

MEMT (*Middle English Medical Texts*, Taavitsainen et al. 2004) contains about half a million words. It covers the period from about 1375 to 1500, and contains 86 documents representing three traditions. We use twelve learned scholastic texts (out of the total 86 MEMT texts) for the document classification experiment (Sections 3.2 and 4.1).

EMEMT (*Early Modern English Medical Texts*, Taavitsainen et al. 2010) contains ca. 2 million words, covering the period from 1500–1700. It contains seven subcategories, including scientific journals, surgical and anatomical treatises, and recipe collections. We use the entire corpus for our experiments with topic modelling (Sections 3.3 and 4.2) and conceptual maps (Sections 3.4 and 4.3), and 15 typical scholastic documents for the document classification experiment (Sections 3.2 and 4.1).

LMEMT (*Late Modern English Medical Texts*, Taavitsainen et al. 2019) also contains about 2 million words, and it covers the period from 1700–1800. We use the entire corpus for the topic modelling experiments (Sections 3.3 and 4.2).

3.2 ARCHER Medical

ARCHER (Biber et al. 1994) is a corpus compiled for studies into register variation and diachronic change in British and American English. It covers the period from 1600 to 1999. It is balanced in terms of genres, regional variety (British and US) and century. Texts from the earlier centuries are available with modernised spelling (Schneider et al. 2017).

---

2. http://www.helsinki.fi/varieng/CoRD/corpora/CEEM/index.html

The size of ARCHER is 3.2 million words, but here we only use the medical part, which contains 214,000 words, in 127 documents. We split ARCHER Medical into two partitions: early, from 1600 to 1800, and late, from 1800 to 2000.

3.3   HIMERA

HIMERA (HIStory of MEdicine coRpus Annotation, Thompson et al. 2016)[3] is a corpus of 35 selected documents from the *British Medical Journal* and four long reports by the London area Medical Officer of Health. It covers the period from the 1850s to the 1960s and has been compiled with the aim of facilitating the semantic exploration of change over time, topics, concepts and language use.

3.4   PubMed Excerpt

While finding suitable documents from the earlier periods is difficult, the situation is reversed after about 1960: there is an abundance of available medical articles, particularly in the online collection PubMed,[4] which contains over 30 million documents. The collection is temporally skewed, reflecting the exponential growth in publication activities. Figure 1 shows how the number of articles published annually increases between 1914 and 2014.
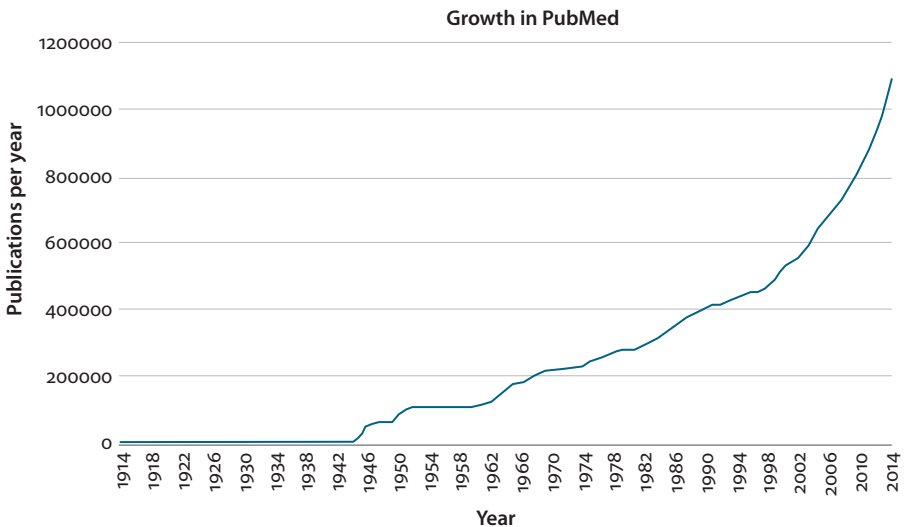

Figure 1. Publication per year in PubMed (Funk 2015)

3. http://www.nactem.ac.uk/himera/

4. https://pubmed.ncbi.nlm.nih.gov

Compared to our other corpora, PubMed is different and possibly problematic in at least three respects:

1. It is a monitor corpus, which leads to a temporal skew, i.e. there are far more articles in the later periods.
2. The database does not distinguish articles based on their importance or typicality. The sample includes articles that potentially address a larger non-expert audience.
3. In most cases, only the abstracts are freely available for download.

Faced with the challenge to create a representative sampler of PubMed, we chose the following workaround: we use a corpus of Spanish-English medical translations. The MeSpEN corpus (*English-Spanish Medical Machine Translation and Terminologies*, Villegas et al. 2018) contains 127,619 PubMed abstracts totalling about 20 million words.[5] As only articles that are of public interest are typically translated, this workaround alleviates problem (2).

Even with a more focused sample based on MeSpEN, the problem of temporal skew (1) remains, but this skew is less extreme than in the full PubMed. As practical measure, we capped at 150 abstracts per year for our sampler. The earliest abstract in sampler corpus is from 1964, and there are 150 abstracts per year after 1975. The total size of the sample is one million words.

3.5   Overview of the complete data of our investigation

For our experiments with document classification, we used twelve scholastic documents from MEMT, and 15 scholastic documents from EMEMT as described in Taavitsainen and Schneider (2018). To this we added ARCHER Medical split into two periods, and HIMERA.

Table 1 shows the period of each corpus, the number of documents and words that each corpus contributes, in total and per century. The proportions per century are fairly balanced except for the nineteenth century. This shortcoming is partly alleviated by using pseudo-documents, as described in Section 4.1.

---

5. http://temu.bsc.es/mespen/

Table 1. Contribution of the various corpora by century, in words

	EMEMT	LMEMT	ARCHER Medical	HIMERA	PubMed-ES-EN	Σ
Period	1500–1700	1700–1800	1600–1999	1850–1960	1964–2018	1500–2018
Documents	463	166	127	39	6,622	7,417
1500s	772,204	–	–	–	–	772,204
1600s	1,378,519	–	24,411	–	–	1,402,930
1700s	–	1,577,540	68,081	–	–	1,645,621
1800s	–	8,550	61,700	49,717	–	119,967
1900s	–	–	65,829	30,337	1,063,054	1,159,220
Σ	2,150,723	1,586,090	220,021	80,054	1,063,054	5,099,942

3.6   Limitations of the data

Using large datasets helps to overcome sparse data issues (cf. Michel et al. 2011): the standard error shrinks in relation to the square root of the sample size, and asymptotically converges to zero. It does not help, however, against a principled bias. Keller and Lapata (2003) and Lapata and Keller (2005) show that, for some tasks, particularly for obtaining word bigram frequencies, larger is typically better: comparing larger web corpora to the balanced 100 million-word BNC shows that large web corpora, even if less balanced, typically deliver better correlations to human plausibility judgements for word bigram sequences. Yet for other tasks, for example the attachment of prepositional phrases, BNC performs better than large web-derived corpora.

From the perspective of data sparseness, our approach of collecting whatever we could obtain is the only possible option: our total of 4.8 million words is still much smaller than BNC. But from the perspective of data bias, the comparability of our sources raises some questions and requires careful interpretation. While CEEM, ARCHER and HIMERA have been compiled to be comparable, that is not the case for PubMed. HIMERA may also be biased towards public health. Moreover, our proxy of using translated texts may not necessarily fully reflect the importance of texts. This is potentially a serious shortcoming of our approach.

To assess the influence of *public* health and *public* interest generally, we checked the frequency of all bigrams with *public* as the first word (*public **). While it is absent in EMEMT, it occurs 3 times in LMEMT. In the eighteenth century, before the HIMERA period, we have 14 occurrences (61 per 1 million words), rising only slightly to 24 (70 per 1 million words) in the nineteenth century, the first century to include HIMERA and thus reports by the Officer of Public Health. This suggests that the rise in public interests may have happened before the HIMERA period, and that the inclusion of Medical Officers of Health reports mirrors a general, social

trend. Despite these potential shortcomings, we decided to proceed, but we are aware that our sample is a reduction (cf. Leech 2007).

A strong imbalance between the corpora comes from the slightly different genres: CEEM, for example, contains many recipes, while PubMed contains many clinical studies. The ensuing genre drift can be seen as a curse or as a blessing, as unwanted noise or as a signal of changing genres. We think that genre drift is an inherent element of the change that we want to detect and describe.

4. Methods

After preparing the data (Section 4.1), we apply several state-of-the-art methods from the domain of machine learning. For an introductory overview of methods, see e.g. Schwartz and Ungar (2015) and Grimmer and Stewart (2013).

4.1 Data preparation

Data preparation involves converting the documents from XML or proprietary formats to a tab-separated format, where the first column contains the meta-data, and the second column a document or a part of it, for which we will use the term "pseudo-document". The spelling of documents published before 1850 has been normalised using VARD2 (Baron and Rayson 2008; for details, see Schneider et al. 2017).

The machine learning approaches that we use are sensitive to document length in two opposing ways. Tang et al. (2014) explains the ensuing trade-off: if documents are very short, they only contain very few of the features that are needed by the algorithm, for example content words. But, at the opposing end, if we only use very few very long documents, we obtain only few data points, which means that the accuracy may stay low. Additionally, in long documents the topic of the discourse is likely to evolve, making it difficult to detect topics from overly long documents.

While our data does not contain very short documents,[6] it contains long documents, so that splitting long documents is beneficial. Our heuristic settings for the length of suitable pseudo-documents are impressionistic, based on pilot runs without thorough evaluation. For the experiments with topic modelling (Sections 3.3 and 4.2) and conceptual maps (Sections 3.4 and 4.3), we have cut long documents at every full stop after 200 words in the EMEMT and LMEMT, after 80 words in ARCHER Medical, and for the relatively small HIMERA corpus (and also because we have little data from the nineteenth century) every sentence is a pseudo-document. As the PubMed abstracts are shorter than the full

---

6. The documents are shortest in PubMed (average length 160.5 words).

articles, we have not split them. Table 2 shows the composition by century and pseudo-documents for topic modelling.

Table 2. Contribution of the various corpora by century, in pseudo-documents

	EMEMT	LMEMT	ARCHER-Medical	HIMERA	PubMed-ES-EN	Σ
1500	3,360	–	–	–	–	3,360
1600	6,109	–	292	–	–	6,401
1700	–	7,031	712	–	–	7,743
1800	–	39	702	1,770	–	2,511
1900	–	–	950	1,146	6,622	8,718
Σ	9,469	7,070	2,656	2,916	6,622	28,733

For document classification, we use longer pseudo-documents (up to 2,000 words per pseudo-document, 340 in total), as we are not discriminating by topics, but rather by century. While topics may change over the discourse of the document, the century of composition does not.

4.2   Supervised document classification

Document Classification is often used in Computational Linguistics and media content analysis (Grimmer and Stewart 2013). Almost every task can be formulated as a document classification task. Some of its best-known applications in computational linguistics are:

–   Binary classes of relevant or irrelevant documents for an Information Retrieval task (see Jurafsky and Martin 2009: Chapter 23.1)
–   Authorship identification in forensic linguistics (see Oakes 2014)
–   Positive or negative assessment of a political issue in automated content analysis (see Grimmer and Stewart 2013)
–   Binary class of *spam* or *nospam* in most email clients

In our application, we are going to predict the century from which the document originates in order to address RQ1, i.e. to see which words or constructions are most typical and overused in which century. In the earliest period, we combine two centuries, to obtain a larger class and thus a more balanced class distribution. There are five classes in total. Document classification is a so-called supervised approach: annotated classes (e.g. the time period) are predicted based on the words in the documents.

In typical machine learning scenarios, including ours, the discriminators between the classes are the content words, i.e. all words excluding stopwords. The sequence or syntactic context of the words is typically not considered, so these models are so-called *bag-of-words* models. To include contextual information, sequences of words, so-called *bigrams*, are also often used. Every word or bigram type (above a certain minimal frequency, e.g. $f > 4$) is a feature. This entails that document classification typically uses thousands of features, and more if bigrams are used. Each feature in isolation is usually not a good discriminator between the classes, but many "weak" discriminators together achieve very high classification accuracy.

Many of the algorithms used are well known in quantitative linguistics, for example logistic regression, but while quantitative linguists use multivariate analysis with perhaps a dozen features, thousands of features are standard in document classification. Additional characteristics of machine learning scenarios are that the training set (from which the probabilities are learned) are separated from the application set (on which the accuracy of the algorithm is measured). To alleviate the impact of overadaptation, the training step is often performed several times; so-called cross-validation systematically leaves out one part of the data, and averages over the different trained models. We also use further smoothing techniques, for example the so-called Lasso regularisation, following best practices in machine learning (Schreiber-Gregory 2018: 6).

The features, e.g. words with the highest weights, are the best discriminators between classes, and thus meaningful patterns that we want to detect and interpret as keywords. Perhaps surprisingly, computational linguists and the machine learning community are typically more interested in the choice of algorithm and feature engineering than in interpreting the features themselves.

In technical terms, we are performing keyword extraction by means of document classification. While document classification with logistic regression using keywords is frequently employed, our "opposite" approach, using feature weights as the keyword extraction algorithm, is rare (see e.g. Schneider 2018).

We rely on the tool LightSide[7] for our supervised document classification experiments (Section 5.1). This tool automatically performs cross-validation, evaluation and POS tagging with the Penn Treebank tagset. We use the algorithm of logistic regression as it combines good accuracy and popularity among linguists. We also us Lasso regularisation and 5-fold cross-validation, as it is automatically calculated in the LightSide tool.

---

7.   http://ankara.lti.cs.cmu.edu/side/

4.3   Unsupervised topic modelling

Interpreting a feature space of thousands of items can be demanding, and for a semantic investigation, the fact that synonyms and related words are not considered is a disadvantage. To detect collocations, quantitative linguists exploit the so-called Firthian hypothesis (Firth 1957). Sahlgren (2006) shows that the size of context in Firthian approaches plays a crucial role: using adjacent words (or very small observation windows) delivers classical collocations, i.e. syntagmatic relations, while expanding the window to include increasingly more context delivers results on the paradigmatic axis, i.e. synonyms, antonyms, associations and related words. This fact is exploited by distributional semantics (e.g. Baroni and Lenci 2010).

The insights into the relationship between words, discourses and contexts has given rise to a usage-based perspective on language (Bybee 2007, Janda 2013), as well as to several distributional methods including topic modelling (Blei 2012).[8] Topic modelling combines document classification with keyword generation, based on the strong semantic unity of the discourse of a topic and of a document. Following Church (2000), the chance of a word to occur in a document radically increases if the same word has occurred before in the same document. While a given word w generally has probability p(w), the chance of seeing w a second time, if it has occurred once, is much closer to p(w)/2 than p(w)^2, which we would expect if the first and second appearance were independent.

The enormous increase of probability of previously seen words also applies at the discourse level of topics. Topic models increase the context beyond the level of the document to the topic level.

Topic modelling optimises the Bayesian probability

p($topic|document$) * p($word|topic$)

for all given documents in a collection. It thus combines document classification (p($topic|document$)) and keyword generation (p($word|topic$)). *Documents* and *words* are given, conceptual *topics* are fitted iteratively starting from a random configuration.

The only parameter users need to set is the number of topics. No-pre-annotated data is needed. As the approach is completely driven by the data, it is a so-called unsupervised approach. Unsupervised approaches are theory-agnostic, but this also entails some disadvantages: it is not straightforward to include the external knowledge, the user has relatively little control over the process, large amounts of data are needed, and evaluation is difficult. For our experiments, we use the tool *Mallet*.[9]

---

8.   See also Menzel (this volume), who uses probabilistic topic modelling to identify topics from the Late Modern English part of the *Royal Society Corpus*.

9.   *Machine Learning for Language Toolkit*, see http://mallet.cs.umass.edu/topics.php

4.4   Unsupervised Conceptual Maps with Kernel Density Estimation

Kernel Density Estimation is also an unsupervised distributional semantic method. It calculates semantic distances between words by exploiting the observation that semantically related words, for example *horses, saddle* and *ride*, appear mostly in the same locus, both in the real world and in large corpora. If we calculate mutual co-occurrences across documents, we obtain a data-driven measure of similarity.

Kernel Density Estimates are Kernel functions of these mutual co-occurrences learned from corpus data. An approximating Kernel function, instead of the lexical overlap, is used to obtain smoother results, glossing over data fluctuations. From the co-occurrence matrix, conceptual maps are calculated. These maps show closely related concepts clustering together, while less closely related words appear increasingly further away on the map. Similar words are plotted close to each other on the emerging concept map. Words found in most discourses appear near the centre of the map, while those restricted to very few topics appear on the fringes of the network. Such conceptual maps are a new and not yet widely used visualisation technique. We believe they are intuitive, inspiring and attractive to interpret, and thus a viable method to produce mind maps and sematic networks automatically. We use David McClure's tool *textplot* for calculating the kernel densities, and *Gephi* for displaying the conceptual maps.[10]

5.   Results

5.1   Results of supervised document classification

We classify the texts into five classes: (1) early scholastic documents from MEMT from 1375 to 1500; (2) late scholastic documents from the EMEMT corpus, from 1550 to 1650;[11] (3) ARCHER Medical from 1700 to 1800; (4) ARCHER Medical and HIMERA from 1800 to 1900; and (5) HIMERA from 1900 to 1960. The corpus for these pilot experiments is smaller than for the later experiments.

The classification accuracy on the 340 pseudo-documents is 95.2%, and the Kappa value (improvement over a random choice baseline) is 94.0%. Table 3 gives the so-called confusion matrix, which compares predicted classes in the columns and actual classes in the rows. The highlighted cells in the diagonal give the number of correctly classified documents.

---

10. http://dclure.org/tutorials/textplot-refresh/

11. The documents from EMEMT and MEMT are the same as in Taavitsainen and Schneider (2018).

Table 3. Confusion matrix of the document classification experiment

Act\Pred	EARLY	LATE	17xx	18xx	19xx
EARLY	37	2	0	0	0
LATE	0	75	0	0	1
17xx	0	0	65	1	0
18xx	0	0	2	75	3
19xx	0	0	0	7	71

The top 50 features of the twentieth century (19xx) are shown in Table 4. As there are more than 6000 features in total, the top 50 features only show the tip of the iceberg.

Table 4. The top 50 features of the twentieth century (19xx)

Rank	Feature	Freq.	Feature influence	Rank	Feature	Freq.	Feature influence
1	CD_-RRB-	53	8.751	26	NNS_JJ	49	4.127
2	associated	36	6.903	27	Failure	20	4.057
3	1	56	6.829	28	Suggested	27	4.044
4	shown	39	6.387	29	Type	17	4.032
5	present	47	6.005	30	Patients	54	3.995
6	2	50	5.672	31	Cancer	11	3.979
7	recent	20	5.539	32	Between	54	3.949
8	obtained	40	5.480	33	Results	40	3.937
9	group	33	5.249	34	Infection	19	3.934
10	3	46	5.230	35	Been	65	3.853
11	-LRB-_NN	56	5.210	36	NN_-LRB-	61	3.823
12	4	38	5.010	37	NNP_CC	37	3.816
13	NN_WDT	59	4.919	38	RRB-_.	49	3.811
14	VBN_VBN	61	4.913	39	VBP_VBN	65	3.800
15	NNS_-LRB-	50	4.840	40	LS_-RRB-	25	3.777
16	CC_NNP	44	4.799	41	NNS_VBP	62	3.767
17	per	51	4.792	42	Cent	41	3.752
18	fact	23	4.459	43	Direct	12	3.751
19	due	36	4.364	44	Areas	11	3.749
20	<CLOSEPAREN>	70	4.346	45	experimental	13	3.716
21	<OPENPAREN>	70	4.346	46	Normal	41	3.714
22	._JJ	52	4.301	47	CD_,	56	3.692
23	-LRB-_NNP	28	4.232	48	Clinical	32	3.670
24	FW_,	25	4.224	49	._EX	32	3.645
25	IN_JJS	34	4.143	50	increasing	17	3.626

The features that are given in boldprint can be easily interpreted. The top 10 features can be interpreted as follows:

1. The tag sequence **CD-RRB-** stands for cardinal number followed by a right, i.e. closing bracket, matching the ending of references or additional numerical information in brackets. This sequence is the strongest feature of the twentieth century.
2. In a data-science perspective, detection and measurement of correlations are very important, which explains **associated**.
3. Numbers and lists are more frequent than in any other century. Also, there is a tendency to use cardinal numbers also for small numbers.
4. The main task of present-day publications is to show *results* (rank 33). In addition, passive forms (rank 14) and the resultative past present (rank 39) are further increasing, and also reduced relative clauses, e.g. *results shown* (see Hundt et al. 2012).
5. Authors **present** their findings and interpret the **present** data.
6. See point 3.
7. Referencing to and advancing on **recent** literature and addressing **recent** trends or diseases are a top priority in present-day research.
8. Results are **obtained** by empirical, often clinical studies.
9. Clinical studies divide the patients into **groups**. Results involve differences *between* (rank 32) them in *clinical* (rank 48) studies.
10. See point 3.

Not every feature has an immediately obvious interpretation (see e.g. 13 and 22), but most can be determined by qualitative analysis.

For space reasons, a list of selected strong positive lexical and POS-based features and short explanations following an arrow (→) is given in Table 5. *Lexical features* mostly point to differences in content, *POS-based* features often also to style.

In sum, most of the strongest features can be linked with their characteristic textual functions. For instance, a search for features of public health reveals that the nineteenth century was a key period, with the sanitary revolution that started in England in the 1830s and 1840s (e.g. Scally 2014).

Although document classification has been accused of being a "black box" (e.g. Broersma and Harbers 2018), such fears can be alleviated by an inspection and interpretation of the features, and an evaluation of the process.

It is worth mentioning that also many features ranking lower than the top 50 are still informative. For example in the twentieth century, **VBP_VBG** (ranking 277) indicates a general increase in progressive forms, and key terms such as *data* at position 74, *number* at 81, and *research* at position 118, illustrate the importance of quantitative approaches and research generally.

Table 5. Document classification features and our suggested interpretation

Group	Lexical features	POS-based features
Early Scholastic (EARLY) *	*sickness, evil* (→ religion)	DT_DT (*each an ounce →* recipes)
	vein, face, (*take*) *heed* (→ moral)	*called* NNP (→ definitions, scholastic style)
	wit, wise (→ belief in authorities)	PRP_CC (→ narrative style)
	swelling, know, cold, humours (→ humours)	
	liver, purgation, vicious, melancholy (→ humours)	
	wicked (→ moral)	
Late Scholastic (LATE) **	*because* (→ argumentation)	FW_FW (→ Latin terms)
	shall (→ moral)	<AMPERSAND> (→ narrative style)
	Galen, unto, says, learned, being, reason (→ scholasticism)	PRP MD (*we shall →* obligation, exhortative style)
	called (→ definitions)	
	according (→ belief in authorities)	
	philosopher, Aristotle, thereby (→ linguistic change)	
	thereof (→ linguistic change)	
	humour, must (→ moral)	
	utter, writes (→ belief in authorities)	
	furthermore (→ argumentation)	
	Avicenna, Greek, whom, beasts, virtue (→ moral)	
17XX (17)	*sometimes, several* (→ variations, exceptions)	._CC (→ narrative style)
	colour (→ *experiments, observations*)	JJR_NNS (comparisons)
	applied, account, observed, method (→ onset of empirical research)	
	effects (→ onset of empirical research)	
	frequently (→ onset of empirical research)	
	found (→ onset of empirical research)	
	than (→ comparing groups and individuals)	
	informed, therefore (→ argumentation)	
	quantities (→ dosage)	
	prevent (→ public health)	
	whence, degree (→ dosage)	
	gentleman (→ society)	
	people (→ society)	
	however (→ argumentation, allowing contradicting facts)	
	necessary (→ obligation)	

Table 5. (*continued*)

Group	Lexical features	POS-based features
18XX (18) [†]	*extent* (→ dosage, quantitative research)	FW_FW (→ Latin terms),
	affections, character (→ moral)	<AMPERSAND> (→ narrative style)
	treatment (→ public health)	PRP MD (*we shall* → obligation, exhortative style)
	hospital (→ public health)	
	system (→ systematic research),	
	regard, fibres, journal (→ publication)	
	pathology (→ systematic research)	
	patient (→ public health)	
	attack, disease, sensations, fatal (→ public health)	
19XX (19) [‡]	*associated, shown, present, obtained, recent, group, fact, due, failure, suggested, type, patients, cancer, results, infection, been, recently, described* (→ resultative, passive)	VBN_VBN (perfect passive, *these disorders have been referred to* → linguistic change)
		JJ_JJ (complex terms)
	experimental, clinical, increasing, chronic, rate, feature, study, important, data, number (→ Data Science)	NN_NN (noun compounds
	work, research, scientific, because, seems, findings, typical, significant, conclusion (→ style, schema with little variation)	VBP_VBN (passive forms)
		NNS_VBN (reduced relative clauses)

*, ** These features are in line with Taavitsainen and Schneider (2018).

† Taavitsainen et al. (2019) point out how public health and professionalised medicine is a determinant characteristic of the eighteenth century, the main period of the LMEMT corpus, and this trend increases further in the nineteenth century.

‡ For the top 50 features of the twentieth century, see Table 4.

5.2   Results of unsupervised topic modelling

While document classification offers detailed insights, sifting long lists is laborious, and it is difficult to gain an overview. Only human readers know that features like *group* (rank 9) and *patients* (rank 30) are closely related because *clinical* (rank 48) studies investigate groups of patients, or that *results* (rank 33) and *present* (rank 5) are related as studies present results.[12] Such semantically coherent topics could

---

12. As we did not POS-tag our texts, some of the instances of *present* are likely to be adjectives, expressing a deictic function.

give one an overview of the data, but they are not delivered automatically, and it is difficult to create them systematically.

This is where topic modelling may be able to help, as it extracts topics automatically from the context. Users only need to set the number of topics. Table 6 gives the results of a run with 10 topics.

Table 6. Output of a topic model from Mallet. Raw output (except for boldprint) in column 5, strongest period in column 2, and our manual interpretation in column 1

Manually suggested label	Strongest in corpus	Topic ID	Weight	Keywords
Body Parts	LMEMT	0	0.08522	*blood body parts air water nature quantity matter part animal heat stomach spirits great motion vessels salt spirit bodies*
Clinical Study: Risk	PubMed	1	0.04623	*patients study health results years age risk care conclusions methods population clinical children factors hospital group objective high analysis*
Professional Practice	LMEMT	2	0.10215	*physicians great nature practice physician diseases time made dr medicine medicines knowledge men method make general art good persons*
Scholastic	EMEMT	3	0.07441	*man hath things men body nature great good time god thing doe doth reason make life moon called made*
Clinical Study: Diagnosis	PubMed	4	0.06087	*patients cases clinical treatment disease diagnosis case patient years results surgery study group surgical present months performed complications renal*
Blood & Wounds	ARCHERM EARLY	5	0.0599	*part parts made wound bones blood side head de bone vessels left eye veins called arteries small uterus heart*
Humours	EMEMT	6	0.06337	*good body hot cold doth called blood man humours stomach hath dry heat water great moist flesh urine liver*
Clinical Study: Significance	PubMed	7	0.03628	*patients levels group results studied mg found blood cells study serum activity significant obtained samples normal values ml showed*
Disease & Symptoms	ARCHERM EARLY	8	0.10982	*time patient day disease pain great days child fever years symptoms case found water blood hours cold cure head*
Recipes	EMEMT	9	0.04982	*water put half oil ounce powder make wine ounces drink white good ii ana morning till juice leaves pound*

The topics suggested by Mallet look convincing and consistent, but the fact that three topics (IDs 1, 4, and 7) are very closely related is unsatisfactory. There is typically a trade-off between setting too few topics, which has the effect that important topics are missed (limited recall), and too many topics, which leads to a situation in which some topics largely overlap (limited precision).

It is hard to validate topic models, as no uncontested evaluation metric exists, unlike in supervised approaches, where accuracy of the prediction is part of the model. Röder et al. (2015) suggest using distributional semantics to evaluate the consistency of topics: the more closely related the words in a cluster in terms of semantic similarity are, the better the topic model is judged to be. While such an evaluation covers a part of the task, it is also problematic: distributional semantics is a related approach, which entails a danger of circularity. Moreover, lexical semantics does not necessarily capture associations and entailment, which are also important. In addition to synonymy, distributional semantics can also cover association "a notion of relatedness that is much looser than that captured in formal synonymy" (Fitzmaurice et al. 2017: 25). In practice, however, topic models bring up more discursive association than distributional semantics, because they have a strong component of discourse unity due to the document classification component of the algorithm (see Section 4.3). For example, the fact that recipes may involve a *powder put* into *wine* to *drink* is not known in a semantic system based on synonymity.
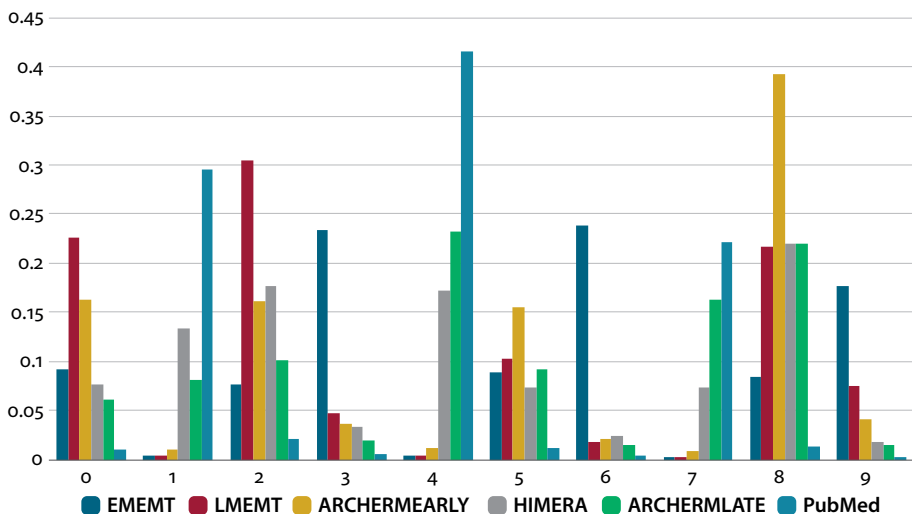

Figure 2. Distribution of topics across corpora representing the five periods of this study

After interpreting the topics and finding them sufficient, we can compare their distribution across time. Table 6 shows the most important period for each topic, and Figure 2 offers a relative comparison of the contribution of the topics per period. We can glean the following insights from Figure 2:

– Body parts (topic 0) are most important in LMEMT, while humours and scholastic thought were even more important in EMEMT.
– Professional *practice* (topic 2) and the *knowledge* of doctors (*dr*), gained particular importance in the LMEMT period, whereas in the later periods they are largely taken for granted.
– Scholastic thinking (topic 3) loses importance more slowly than the belief in the four humours (topic 6). Recipes (topic 9) also disappear more slowly.
– The PubMed period almost exclusively contributes to the clinical study topics (1, 4, 7), retreating to technical jargon. Even obviously central aspects of medicine, such as body parts (topic 0) and symptoms (topic 8) are inherently assumed instead of explicitly mentioned.
– From this perspective, the move from symptoms and specific diseases (topic 8) to diagnostics (topic 4) can be seen as a further step of professionalisation. Topic 4 is the most peaked topic, dominated by the PubMed period.
– Although topic 1 (clinical study: *risk*) is strongest in the PubMed period, its mean is considerably earlier than topics 4 and 7. The keywords reveal that public *health, health care* and *care* for the *population* are also in the foreground in this topic. This distinguishes topic 1 reasonably well from 4 and 7.
– Topic 2 (professional practice) and topic 5 (blood and wounds) are the most stable topics: they are important at each period, along with Topic 0 (body parts) and 8 (disease and symptoms). These topics are still largely implicit in PubMed as discussed above.

5.3 Results of Unsupervised Conceptual Maps with Kernel Density Estimation

A disadvantage of document classification is that the lists of words that they report are very long. Topic modelling helps us here by partly bridging the gap from words to conceptual topics. But *between* the topics we cannot see how close they are to each other, and *within* the topics we cannot see how closely the keywords are related. A better overview may be obtained using a method which keeps all keywords but arranges them in a more interpretable order than a 1-dimensional list may.

Conceptual maps deliver a 2-dimensional view, placing words not only in relation to the class to be predicted (document classification) but also in relation to

each other, and possibly metadata – we are including the corpus name representing the period.

The conceptual maps created with Kernel Density Estimation (Section 4.4) are complex, and a complete map cannot be displayed in detail on a single page. Instead of displaying a complete map in which the keywords are not legible we display four excerpts in Figures 3–6. The corpus name is highlighted in yellow. We use the 400 most frequent words (excluding function words, using the default stopword list of *textplot*). The map is arranged largely chronologically (from right to left), and the period defines its shape. The fact that the temporal dimension turns out to be the dominant axis is a first indication of success.

Our journey proceeds in a diachronic order, starting with Figure 3, which is at the far right of the complete map. We see that *virtue* is closest to the EMEMT period, but also indications of the four humours (*moist, dry, cold, fire*). Digestion is

Figure 3. EMEMT and the genre of recipes

an important topic (*stomach, belly, nourishment*). On the top we can see the recipe corner with many keywords (e.g. *mixed, powder, boil, distilled*).

Figure 4 further reveals that *humours* and *God* are key terms of the period, as are the *books* of the learned scholastics such as *Galen* (see Taavitsainen et al. 2019). Some terms of scholastic thinking, particularly the desire to define (*called*) and classify (*sort*) are drawn into the centre of the map. Other scholastic concepts, e.g. *reason, words, learned* and *mind* all appear in the middle between EMEMT and LMEMT, indicating their importance in both periods.

The centre of Figure 4 is also the centre of the conceptual map. It contains words that are important in all periods of the medical discourse (*medicine, remedy, bone, rest*), other body *parts* (*eyes, mouth, head)* and generally frequent words (*time, make, made, great*). Semantically closely related words often appear close together (e.g. *long, beginning, end*). The term *txt* is the file extension which we explicitly left in as a proxy to the centre. The map was plotted with the option of avoiding overlapping labels, which has the effect that the centre gets artificially stretched, like a city centre inlet in a town map. The map is not true to scale in the centre: the distance between *made* and *make* appears larger than what the underlying numerical data suggests.

Leaving the centre towards the top brings us to the surgical area (*surgeon, cut, swelling, breast, open*, and further off *wound*), although this topic area is less clearly defined as others. Further to the left, we see that *inflammation, skin* and *pain* are related, and pain is also closely related to childbirth (*child, woman*). Moving to LMEMT at the left bottom shows us that *people*'s constitution and the importance of *exercise* is increasing. The keywords *prove, respect*, and *account* foreshadow the ever-increasing shift towards empirical science.

Figure 5 takes us from LMEMT to the late ARCHER period via HIMERA. The way is paved by words increasingly associated with empirical science: from *quality* to *degree, observed, effects, evidence* based on *numbers*, subtle *differences* and *increases. Symptoms* is also at the centre of the part of the map (see topic 8 in Section 4.2). Traces of topic 2 (*practice, knowledge, dr*) can also be discerned. Public health and human care are also topics here, not least because the extremely frequent word *people* (see Figure 4) is pulled far into this direction.

Finally, Figure 6 continues the journey of empirical science, where *evaluated data* and creating the right *sample* (as in corpus linguistics) is key, *method* and *material* grow even more important and *analysis* of *factors* is centre-stage, *significant* differences between *test group*s in a *clinical study* are *determined* and *reported*. The map also mirrors the present-day conventions of the genre: *material, results, method* are used frequently, in the text and as titles. The fact that the keywords are
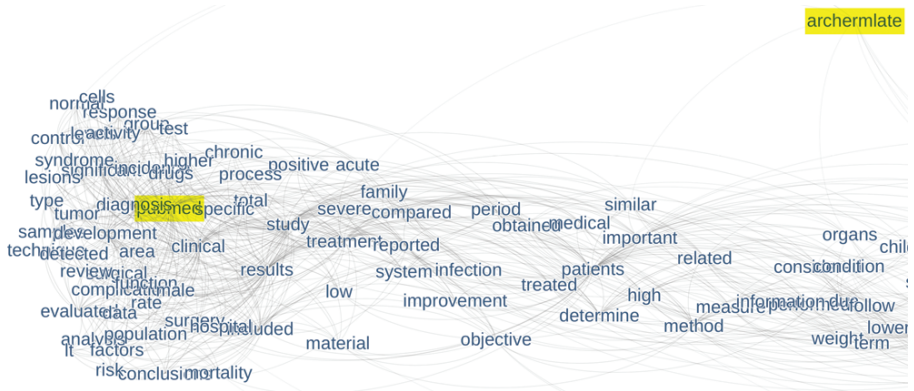
Figure 4. From EMEMT to LMEMT

Figure 5. LMEMT, symptoms and evidence to late ARCHER

Figure 6. From late ARCHER to PubMed and clinical studies

closer together than in any other period indicates that the genre is very narrow. Articles follow trodden patterns (cf. Erman and Warren 2000) and use a restricted vocabulary, except for technical terms such as names of drugs, diseases, proteins and genes, but only the most generic of these (*tumor, infection*) make it to the list of the 400 most frequent words. The fact the PubMed sample only contains abstracts may also contribute to the small distances between keywords.

Surprisingly, keywords referring to experiments only appear quite late: *number, evidence* and *caused* are closest to HIMERA, *methods* is placed between later ARCHER and PubMed, *material* even closer to PubMed; *experimental* is only a keyword in the twentieth century (see Table 4). This finding is in agreement with Bazerman's (1988) study on the *Philosophical Transactions*, showing that actual experiments were few in number up to 1800, and they did not aim to test claims or be reproducible, nor did they use clearly described methods.

Equally surprising is the fact that many obvious keywords of empirical approaches (e.g. *observed, evidence, number*) appear closest to earlier periods. This is not a sign of retreat, but it indicates, firstly, that they were already important at earlier periods and that their early appearance is due to observations becoming important before the rise of the experiment (notice the very long distances between *observed, caused* and *method* in Figures 5–6). Secondly, empirical methods are now tacitly assumed rather than explicitly mentioned: data consists of numbers, significant results entail differences, factor analysis entails statistical regression etc. The keyword *objective*, which at first glance looks like a good candidate for a meaningful keyword, actually represents the search for objective truth only in a minority of cases, whereas the bulk of occurrences are due to the fixed phrase *the objective of our study is*, often followed by *to determine*. Methods for word sense disambiguation, or building maps of concepts could improve the accuracy of the results.

6.   Conclusion and future prospects

We have reached the end of our journey across over 500 years of medical discourse. We have explored the research question of how well data-driven approaches work for the analysis of medical discourse, and what patterns we can discover, using document classification, topic modelling and conceptual maps. We have observed a very long time span from a remote bird's-eye perspective in an exploratory fashion. We have applied visualisation techniques to show us how radically the domain has changed across the centuries, and where the foci of the different periods can be found.

We return to our four research questions again in the following:

RQ1.   Can a systematic comparison of quantitative frequencies of terms and linguistic features detect historical developments in the medical domain? Are the reported differences interpretable?

Using document classification, we have seen that most of the top features are straightforward to interpret, and 95% of the documents are classified in the correct century, which indicates that the features are reliable.

RQ2.   Can the underlying topics and concepts be detected by quantitative methods?

While the validation of the results of topic modelling is intrinsically difficult, all the 10 topics identified by *Mallet* are consistent and easily interpretable. That three topics (1, 4 and 7) largely overlap (see Table 5) is less convincing. Nevertheless, we learned that professional practice was already a crucial topic in the LMEMT period, that belief in humoral theory wanes faster than scholastic thinking, that PubMed is dominated by clinical studies and statistics, while descriptions of symptoms (topic 8), which dominated intermediate periods, are backgrounded. While these results are perhaps unsurprising, the chosen method provides a fascinating illustration of how radically the domain has changed across the centuries.

RQ3.   Can the underlying assumptions, the trends in the history of thought, also be uncovered?

The fact that the shape of the overall conceptual map, based on Kernel Density Estimation, is stretched far on the temporal axis, and that chronological order emerges automatically, indicates that period is the dominant semantic parameter. The conceptual maps further enrich the picture gained in the other approaches, identifying corners with *recipes*, *surgery*, and *see*; the stepwise rise of empirical science, from *knowledge* (LMEMT) to *evidence, increase, differences* and *numbers* (ARCHER and HIMERA) to *data, significance factors, tests response* (PubMed).

Conceptual maps also identify PubMed as a highly formulaic genre, which is evident from the fact that the most frequent terms appear close together. But this observation may be due to the PubMed comprising only abstracts. This raises the question of whether we have used the right instruments and data to observe our world? Are we in a situation of the emperor with no clothes (Leech 2007: 134), unable to see ourselves, or that the earth is not flat as we are standing on it? RQ4 comes in here.

RQ4.   How far does our selection of material affect our results? Can we assess its potential impact?

While CEEM, ARCHER, and HIMERA have explicitly been compiled for internal comparability, comparability *across* corpora was not an aim of the compilers. Our results also partly reflect genre drift (which is intentional) and partly corpus-specific characteristics. In particular, the interest in public health in HIMERA, and the fact that our PubMed collection (MeSpEN) consists of abstracts, has probably had an impact on our results. We have cautioned readers where we suspect the results have been influenced by our choice of corpora. While we have seen how frequency of *public** rises already before the HIMERA period, we cannot assess its full impact: neither the benefit of the doubt for the data nor the circularity problems in its compilation can be fully avoided. The compilation of more comparable corpora, continuing the tradition of CEEM, is a desideratum for future research.

Despite these caveats, our exploration has allowed us to trace the development of the medical discourse, from scholastic thinking and empirical approaches, first based on observation and then increasingly on hypothesis-testing, data, and sampling. In a sense, the wheel has come full circle. While data-driven approaches are increasingly complementing hypothesis-driven research in medicine and biomedicine (Ananiadou et al. 2006), our investigation of this field of writing has now used these data-driven approaches. In that respect, the journey has not ended but is at the beginning. As a next step, we plan to build maps of concepts, combining topic models and conceptual maps.

Acknowledgements

References

Ananiadou, Sophia, Douglas B. Kell, and Tsujii, Jun-ichi. 2006. "Text Mining and Its Potential Applications in Systems Biology." *Trends in Biotechnology* 24 (12): 571–579. https://doi.org/10.1016/j.tibtech.2006.10.002

Baron, Alistair, and Paul Rayson. 2008. "VARD 2: A Tool for Dealing with Spelling Variation in Historical Corpora." In *Proceedings of the Postgraduate Conference in Corpus Linguistics*, Aston University, Birmingham, UK, 22 May 2008. http://eprints.lancs.ac.uk/41666/1/BaronRaysonAston2008.pdf

Baroni, Marco, and Alessandro Lenci. 2010. "Distributional Memory: A General Framework for Corpus-based Semantics." *Computational Linguistics* 36 (4): 673–721. https://doi.org/10.1162/coli_a_00016

Bazerman, Charles. 1988. *Shaping Written Knowledge*. Madison: University of Wisconsin.

Biber, Douglas, Edward Finegan, and Dwight Atkinson. 1994. "ARCHER and Its Challenges: Compiling and Exploring a Representative Corpus of Historical English Registers." *Creating and Using English Language Corpora: Papers from the 14th International Conference on English Language Research on Computerized Corpora, Zürich 1994*, ed. by Udo Fries, Peter Schneider, and Gunnel Tottie, 1–13. Amsterdam: Rodopi.

Blei, David. 2012. "Probabilistic Topic Models." *Communications of the ACM* 55 (4): 77–84. https://doi.org/10.1145/2133806.2133826

Broersma, Marcel, and Frank Harbers. 2018. "Exploring Machine Learning to Study the Long-Term Transformation of News." *Digital Journalism* 6 (9): 1150–1164. https://doi.org/10.1080/21670811.2018.1513337

Bybee, Joan. 2007. *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301571.001.0001

Church, Kenneth. 2000. "Empirical Estimates of Adaptation: The chance of Two Noriegas is closer to p/2 than p^2." *Proceedings of the 17th Conference on Computational linguistics (COLING 2000)*, 180–186. Stroudsburg: Association for Computational Linguistics. https://doi.org/10.3115/990820.990847

Conklin, Kathy, and Norbert Schmitt. 2012. "The Processing of Formulaic Language." *Annual Review of Applied Linguistics* 32: 45–61. https://doi.org/10.1017/S0267190512000074

Erman, Britt and Beatrice Warren. 2000. "The Idiom Principle and the Open Choice Principle." *TEXT* 20 (1): 29–62.

Firth, John Rupert. 1957. "A Synopsis of Linguistic Theory 1930–1955." *Studies in Linguistic Analysis* [Special Volume of the Philological Society]: 1–32. Oxford: Blackwell.

Fitzmaurice, Susan, Justyna A. Robinson, Marc Alexander, Iona C. Hine, Seth Mehl, and Fraser Dallachy. 2017. "Linguistic DNA: Investigating Conceptual Change in Early Modern English Discourse." *Studia Neophilologica* 89 (sup1): 21–38. https://doi.org/10.1080/00393274.2017.1333891

Funk, Christopher. 2015. "Concept Recognition and Its Application for Protein Function Prediction." *Computational Biology Thesis Defense*. University of Colorado. https://www.slideshare.net/csfunk/funk-defense-upload

Ghanem, Salma. 1997. "Filling the Tapestry: The Second Level of Agenda Setting." In *Communication and Democracy: Exploring the Intellectual Frontiers in Agenda-Setting Theory*, ed. by Maxwell McCombs, Donald L. Shaw and David Weaver. 3–14. Mahwah, NJ: Lawrence Erlbaum.

Grimmer, Justin, and Brandon Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21 (3): 267–297. https://doi.org/10.1093/pan/mps028

Hilpert, Martin, and Stefan Gries. 2016. "Quantitative Approaches to Diachronic Corpus Linguistics." In *The Cambridge Handbook of English Historical Linguistics*, ed. by Merja Kytö, and Päivi Pahta, 36–53. Cambridge: Cambridge University Press.
https://doi.org/10.1017/CBO9781139600231.003

Hundt, Marianne, David Denison, and Gerold Schneider. 2012. "Relative Complexity in Scientific Discourse." *English Language and Linguistics* 16 (2): 209–240.
https://doi.org/10.1017/S1360674312000032

Janda, Laura A. (ed.) 2013. *Cognitive Linguistics: The Quantitative Turn. The Essential Reader*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110335255

Jurafsky, Daniel, and James H. Martin. 2009. *Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics*. 2nd edition. Upper Saddle River, NJ: Prentice-Hall.

Keller, Frank and Mirella Lapata. 2003. "Using the Web to Obtain Frequencies for Unseen Bigrams". *Computational Linguistics*, 29:3, 459–484. https://doi.org/10.1162/089120103322711604

Lapata, Mirella and Frank Keller. 2005. "Web-based Models for Natural Language Processing". *ACM Transactions on Speech and Language Processing*, 2:1, 1–31.
https://doi.org/10.1145/1075389.1075392

Late Modern English Medical Texts 1700–1800 (LMEMT). 2019. Compiled by Taavitsainen, Irma, Turo Hiltunen, Ville Marttila, Päivi Pahta, Maura Ratia, Carla Suhr and Jukka Tyrkkö. Amsterdam: John Benjamins. CD-ROM published with a book. https://doi.org/10.1075/z.221

Leech, Geoffrey. 2007. "New Resources, or Just Better Old Ones? The Holy Grail of Representativeness." In *Corpus Linguistics and the Web*, ed. by Marianne Hundt, Nadja Nesselhauf, and Carolin Biewer, 133–149. Amsterdam: Rodopi.

Michel, Jean-Baptiste, Shen, Yuan Kui, Aiden, Aviva P., Veres, Adrian, Gray, Matthew K., Pickett, Joseph P., Hoiberg, Dale, Clancy, Dan, Norvig, Peter, Orwant, Jon, Pinker, Steven, Nowak, Martin A. & Aiden, Erez Lieberman. 2011. Quantitative analysis of culture using millions of digitized books. *Science* 331(6014): 176–182.

Oakes, Michael P. 2014. *Literary Detective Work on the Computer*. Amsterdam & Philadelphia, PA: Benjamins. https://doi.org/10.1075/nlp.12

Roberts, Marilyn, Tzong-Horng (Dustin) Dzwo, and Wayne Wanta. 2002. "Agenda Setting and Issue Salience Online." *Communication Research* 29: 452–465.
https://doi.org/10.1177/0093650202029004004

Röder, Michael, Andreas Both, and Alexander Hinneburg. 2015. "Exploring the Space of Topic Coherence Measures." *Proceedings of WSDM'15*, February 2–6, 2015, 399–408, Shanghai, China. https://doi.org/10.1145/2684822.2685324

Sahlgren, Magnus. 2006. The Word-Space Model: Using Distributional Analysis to Represent Syntagmatic and Paradigmatic Relations between Words in High- Dimensional Vector Spaces. PhD dissertation, Stockholm University.

Scally, Gabriel. 2014. "Public Health Profession." In *Encyclopedia of Health Economics, Vol. 3*, ed. by Anthony J. Culyer, 204–209. San Diego: Elsevier.
https://doi.org/10.1016/B978-0-12-375678-7.00402-8

Schneider, Gerold. 2018. "Differences between Swiss High German and German High German via Data-Driven Methods." *Proceedings of the 3rd Swiss Text Analytics Conference* (SwissText 2018), Winterthur, Switzerland, ed. by Mark Cieliebak, Don Tuggener and Fernando Benites, 17–25. http://ceur-ws.org/Vol-2226/

Schneider, Gerold, Eva Pettersson, and Michael Percillier. 2017. "Comparing Rule-Based and SMT-Based Spelling Normalisation for English Historical Texts." *Proceedings of the NoDaLiDa 2017 Workshop on Processing Historical Language*, Gothenburg, Sweden, ed, by Gerlof Bouma and Yvonne Adesam, 40–46.

Schreiber-Gregory, Deanna. 2018. "Regulation Techniques for Multicollinearity: Lasso, Ridge, and Elastic Nets." *Proceedings of Western Users of SAS Software Conferences 2018*, September 5–7, 2018, Sacramento, California. https://www.lexjansen.com/wuss/2018/131_Final_Paper_PDF.pdf

Schwartz, H. Andrew, and Lyle H. Ungar. 2015. "Data-Driven Content Analysis of Social Media: A Systematic Overview of Automated Methods." *The ANNALS of the American Academy of Political and Social Science* 659 (1): 78–94. https://doi.org/10.1177/0002716215569197

Sinclair, John and Ronald Carter. 2004. *Trust the Text: Language, Corpus and Discourse*. London: Routledge. https://doi.org/10.4324/9780203594070

Steinberger, Ralf, Aldo Podavini, Alexandra Balahur, Guillaume Jacquet, Hristo Tanev, Jens Linge, Martin Atkinson, Michele Chinosi, Vanni Zavarella, Yaniv Steiner, and Erik van der Goot. 2015. "Observing Trends in Automated Multilingual Media Analysis." *Proceedings of the Symposium on New Frontiers of Automated Content Analysis in the Social Sciences (ACA'2015), Zürich, Switzerland*, 1–3 July, 1–8. https://arxiv.org/abs/1603.02604

Taavitsainen, Irma, Turo Hiltunen, Anu Lehto, Ville Marttila, Päivi Pahta, Maura Ratia, Carla Suhr and Jukka Tyrkkö. 2019. *Late Modern English Medical Texts: The Corpus*. In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221

Taavitsainen, Irma, Päivi Pahta, Turo Hiltunen, Martti Mäkinen, Ville Marttila, Maura Ratia, Carla Suhr, and Jukka Tyrkkö. 2010. *Early Modern English Medical Texts: Corpus*. In *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen, and Päivi Pahta. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.160

Taavitsainen, Irma, and Gerold Schneider. 2018. "Scholastic Argumentation in Early English Medical Writing and Its Afterlife: New Corpus Evidence." In *From Data to Evidence in English Language Research*, ed. by Carla Suhr, Terttu Nevalainen, and Irma Taavitsainen, 191–221. Leiden: Brill.

Taavitsainen, Irma, Gerold Schneider, and Peter Jones. 2019. "Topics of Eighteenth-Century Medical Writing with Triangulation of Methods: LMEMT and the Underlying Reality." In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen, 31–74. Benjamins: Amsterdam. https://doi.org/10.1075/z.221.03taa

Tang, Jian, Zhaoshi Meng, Xuanlong Nguyen, Qiaozhu Mei, and Ming Zhang. 2014. "Understanding the Limiting Factors of Topic Modeling via Posterior Contraction Analysis." *Proceedings of the 31st International Conference on Machine Learning*, 32(1), ed. by Eric P. Xing, and Tony Jebara, 190–198. http://proceedings.mlr.press/v32/tang14.html

Thompson, Paul, Riza Theresa Batista-Navarro, Georgios Kontonatsios, Jacob Carter, Elizabeth Toon, John McNaught, Carsten Timmermann, Michael Worboys, and Sophia Ananiadou. 2016. "Text Mining the History of Medicine." *PLOS ONE* 11 (1): e0144717. https://doi.org/10.1371/journal.pone.0144717

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam: Benjamins. https://doi.org/10.1075/scl.6

Villegas, Marta, Ander Intxaurrondo, Aitor Gonzalez-Agirre, Montserrat Marimon, and Martin Krallinger. 2018. "The MeSpEN Resource for English-Spanish Medical Machine Translation and Terminologies: Census of Parallel Corpora, Glossaries and Term Translations." In *LREC MultilingualBIO: Multilingual Biomedical Text Processing*, Miyazaki, Japan, ed. by Maite Melero, Martin Krallinger and Aitor Gonzalez-Agirre, 32–39, ELRA. http://lrec-conf.org/workshops/lrec2018/W3/pdf/book_of_proceedings.pdf

Medical discourse in Late Modern English

Insights from a multidisciplinary corpus of scientific journal articles

Katrin Menzel

Saarland University

This chapter demonstrates how the *Royal Society Corpus*, a richly annotated corpus of around 48,000 English scientific journal articles covering more than 330 years, can be used for lexico-grammatical and pragmatic studies that contribute to a broader understanding of the development of medical research articles. The Late Modern English period together with several decades before and after this time frame was a productive period in the medical output of the Royal Society. This chapter addresses typical linguistic features of scientific journal articles from medical and related sciences from this period demonstrating their special status in the context of other traditional and emerging disciplines in the corpus data. Additionally, language usage and text-type conventions of historical medical research articles will be compared to the features of corpus texts on medical topics from Present-day English.

1. Introduction

This chapter illustrates how the *Royal Society Corpus* (RSC, Kermes et al. 2016, Fischer et al. 2020) can be used for sociopragmatic and lexico-grammatical studies that contribute to a broader understanding of the history of medical discourse and the structure and development of medical research articles published in English. The linguistically and pragmatically annotated RSC complements other diachronic resources for investigating the evolution of scientific publications from the medical domain and other disciplines. Various types of metadata provide contextual information that can inform work in historical linguistics, sociolinguistics, history

of science or English for specific and academic purposes. These metadata contain sociopragmatically relevant information related to register variables from Systemic Functional Linguistics, such as *Field, Tenor* and *Mode* (cf. Halliday 1987), which influence the linguistic choices of producers of texts.

The corpus contains scientific journal articles, most of which were taken from the *Philosophical Transactions* (*Phil. Trans.*) and the *Proceedings* (*Proc.*) of the Royal Society of London (RS). The RSC covers more than 330 years of research articles ranging from the early stages of the first scholarly journals published in English in the 1660s (for a relatively narrow circle of readers) to scientific publications with a continuously widening reach from the end of the twentieth century, when printed and electronic editions of journals started to complement each other and English had achieved an internationally prominent position in medical research and specialized communication.

Among the RSC texts are various landmark papers by renowned authors that have impacted the scientific publishing culture in Britain and internationally. *Phil. Trans.* and *Proc.* started as general scientific journals and have never been primarily dedicated to medical research like specialized medical journals. They do, however, contain numerous articles on medical topics written by medical researchers and practitioners as well as many articles on health-related topics from life science disciplines such as anatomy, biology, biochemistry, neuroscience and pharmacology. The corpus serves as a resource to trace the development of various aspects related to medical English and to observe some of the characteristics of medical texts in comparison to texts from other scientific disciplines. The RSC has previously been used to trace the overall linguistic development of scientific English across various disciplines (e.g., Degaetano-Ortlieb and Teich 2016, Degaetano-Ortlieb and Teich 2019). In contrast, for the present study, the RSC is specifically filtered for documents on health-related aspects.

Various case studies, handbooks and teaching materials address linguistic aspects of Present-day English (PDE) medical discourse and health care communication in oral and written contexts. Among such works there are many practice-oriented, instructional guides for English-speaking health care specialists for effective and comprehensible oral communication with laypeople and professionals as well as numerous prescriptive instruction manuals on writing contemporary medical research papers and related documents (e.g., Stuart 2007, Goodman and Edwards 2014, Jenicek 2014, Linares, Daly and Daly 2017, Taylor 2018). Such publications often criticize the use of complicated medical jargon in medical discourse and emphasize the use of plain English with easily processable text and unambiguous terms and structures as main success factors for contemporary medical texts. Various corpus-based studies provide insights into PDE medical discourse, e.g., Ferguson (2001) and Roberts (2006) analysing data from doctor-patient

consultations, Fløttum (2006) discussing rhetorical strategies in medical research articles in an interdisciplinary and cross-linguistic study, or Bellés-Fortuño (2016) comparing popular science articles with scientific medical articles. Furthermore, Panocová (2015, 2017a) focusses on medical English terminology and the transparency of neoclassical medical terms for experts and laypeople in comparison to alternative formations such as eponyms or abbreviations. She also addresses the challenges of translating medical vocabulary in such cases where the productivity of certain word-formation processes and the structure of semantic networks and lexical relationships differ across languages (Panocová 2017b). Large specialized corpora of PDE medical texts have been compiled from electronically available resources, e.g., a corpus of more than 85 million words from cardiology journals with online articles available in HTML format (Teufel and Elhadad 2002) or the *Health Science Corpus* with around 4 million words from research articles published in online journals specialized in medicine, biology, biochemistry or biomedicine (Laso 2009: 46ff). Distinctive, conventionalized features of modern medical discourse in English-speaking communities have developed in relation to scientific progress and historical change in medical understanding.

As appropriate data from earlier periods of language has been relatively scarce, diachronic aspects of English medical discourse have been addressed to a much lesser extent than the features of contemporary medical English. Historical studies on medical discourse for specific time periods often involve manual analysis of individual texts or the analysis of relatively small corpora of manuscripts and early printed books. The history of scientific medical articles is closely connected with the activities and publications of the RS and with specialized medical journals like the *Edinburgh Medical Journal* and *The Lancet*, which started to rapidly circulate advances in medicine and surgery enabling the dissemination of new procedures and standardization processes of medical care and its rhetorical discourse. Individual texts from such journals have been integrated in various diachronic corpora (cf. Biber and Finegan 1994, Atkinson 1992, 1999, Salager-Meyer 1999).

The *Corpus of Early English Medical Writing* (CEEM), a corpus of medical documents from various text categories with around 3.75 million words and 750 texts, has recently been released to enable systematic analyses of evolving register conventions and specific developments in medical English from 1375 to 1800 (cf. Pahta and Taavitsainen 2004, Taavitsainen and Pahta 2000, Taavitsainen 2011: 80ff., Taavitsainen and Hiltunen 2019). Its three subcorpora – *Middle English Medical Texts* (MEMT), *Early Modern English Medical Texts* (EMEMT) and *Late Modern English Medical Texts* (LMEMT) – contain different types of expert-to-expert discourse, e.g., surgical treatises, texts on specific methods or substances as well as scientific journal texts, including several articles from the seventeenth and eighteenth century *Phil. Trans.* Taavitsainen (2015) uses samples of early scientific and medical

periodicals from the CEEM to compare English medical news for professional and lay audiences on the basis of keyword analyses and qualitative assessment. The compilation process of the CEEM involved various steps of manual work. Many CEEM texts were manually keyed in from facsimiles. To identify medical articles among the texts in sources such as *Phil. Trans.*, typically the titles of the articles in the tables of contents of the journals served as the methodological point of departure and then a close, qualitative reading of the texts was carried out. By contrast, the steps in the compilation, preprocessing, linguistic annotation and corpus encoding of the RSC as a large corpus of scientific articles from Early Modern, Late Modern and Present-day English were mostly automatic.

The next section will give an overview on the current release version of the RSC. Section 3 will address aspects relating to identifying and analysing corpus texts with medically related discourse topics and participants. Characteristic features of medical articles demonstrating their special status in the context of other emerging disciplines will be described for Late Modern English (LModE) texts and compared to articles on medical topics from PDE.

2.   The RSC 6.0

The RSC is a diachronic corpus of scientific articles from flagship journals of the Royal Society of London published over a time span of more than 330 years. The latest version of the corpus, *RSC 6.0 Full*, includes ca. 300 million tokens and around 48,000 full texts. The RSC does not represent the full variety of English scientific or medical texts; rather, it complements other synchronic and diachronic scientific and medical corpora. However, other resources often have restricted access and come with limited or no linguistic annotations and metadata. They are not always available in the required format or only support basic types of queries such as searches for specific keywords.

The RSC contains the articles published from 1665 onwards in *Phil. Trans.* and *Proc.* and its more specialized successor journals *Phil. Trans. A* and *B* (after 1887) and *Proc. A* and *B* (after 1905). While the early journals used to cover all major scientific disciplines of the time, the A series are dedicated to the mathematical and physical sciences; and the B series cover the biological sciences. Also included in the RSC are texts from additional, more recently launched RS journals on specific academic subfields that had partly been covered in the earlier *Phil. Trans.* and *Proc.*, such as articles on the history of science from the *Notes and Records* and biographical essays from the *Biographical Memoirs of Fellows of the Royal Society*. As pointed out above, the corpus includes numerous articles on medical topics and more broadly on biological and life sciences and offers the possibility to trace

the development of their linguistic characteristics from the early stages of scholarly journals to contemporary publications. Moreover, as a richly annotated, comprehensive source of academic papers, the RSC may be used to verify linguistic statements made on the basis of smaller samples of historical scientific articles so that more robust conclusions can be reached. Fine-grained analyses of evolving discourses on general and specific medical topics, e.g., childbirth or specific illnesses can also be conducted on the basis of those data.

All LModE corpus texts are available for free download and online query in a CQPweb interface (cf. Hardie 2012) from the CLARIN-D centre at Saarland University under a persistent identifier[1] in the release version *RSC 6.0 Open* (see Table 1). It is a subset of non-copyrighted texts from the *RSC 6.0 Full* (cf. Fischer et al. 2020).

Table 1. Size of the *RSC 6.0 Open*

Time period	Texts	Tokens
1665–1699	1,325	2,582,856
1700–1749	1,686	3,414,795
1750–1799	1,819	6,342,489
1800–1849	2,774	9,112,274
1850–1899	6,754	36,993,412
1900–1920	3162	20,159,911
Total	**17,520**	**78,605,737**

The data may be used, for instance, to analyze the use of individual morphological, lexical and grammatical items, multi-word expressions, phrasal structures and discourse-pragmatic features in medical texts. The *RSC 6.0 Open* is suitable for comparisons with other LModE corpora, many of which are searchable through CQPweb as well.[2] The subcorpus of *Late Modern English Medical Texts* (Taavitsainen and Hiltunen 2019) in the CEEM includes eighteenth-century medical texts. The *Coruña Corpus of English Scientic Writing* (Crespo García and Moskowich 2015) contains scientic but no medical texts from the period between 1700 and 1900. The *Corpus of Late Modern English* (CLMET V3.1, 1710–1920 by De Smet et al. 2015) represents non-specialized LModE usage. The RSC with its peer-reviewed journal articles has the advantage of representing a specific text type of professional science communication across a wide time span. It comes with high accessibility and offers various query and visualization options. It has been enriched with linguistic

---

1. https://hdl.handle.net/21.11119/0000-0004-8E37-F.

2. The *Corpus of Late Modern English Texts* and the Coruña subcorpus with LModE history texts (CHET), for instance, have a CQPweb version under https://fedora.clarin-d.uni-saarland.de/cqpweb/ (registration required).

annotations, e.g., for lemmas and parts of speech. The process of creating the RSC is described in more detail in Kermes et al. (2016), Knappen et al. (2017) and Menzel, Knappen and Teich (2021).

Computational language models have been applied for exploring diachronic aspects of language use in the RSC from a macro- and micro-analytic perspective, e.g., n-gram language models and relative entropy (Kullback-Leibler divergence). Surprisal, an information-theoretic operationalization, has been calculated for various units and annotated in the corpus to measure the probability of a linguistic unit to occur in a given textual context (cf. Fischer et al. 2017). Examples of how surprisal might be used in RSC-based research in scientific and medical fields are provided in Menzel and Degaetano-Ortlieb (2017: 225ff.). Word embeddings are used for the identification of semantically similar and potentially relevant linguistic items as well as for observing semantic change of groups of related lexemes (Fankhauser and Kupietz, n.d.).

Various types of metadata in the RSC provide users with contextual information. They were either extracted from available resources such as JSTOR and RS data or generated during the corpus compilation process with automatic or semi-automatic procedures. Topics derived from probabilistic topic modelling serve as indicators for particular scientific disciplines and scientific versus general language (*Field*). Metadata on authors and other participants in medical discourse provide information on social background and relations (*Tenor*). The metadata include variables such as author affiliation and role (e.g., editor, author or communicator of a paper written by another author) and a unique fellow identification number, by which it is possible to link these data with those contained in other databases such as the *Royal Society Fellows Directory*. Metadata on text categories within academic journals and the role of language, e.g., whether texts were intended to be read aloud at scientific meetings, are related to discourse *Mode*.[3] Each metadata file pertaining to the individual RSC texts contains systematically structured contextual information relevant for linguistic analyses with a socio-historical and socio-pragmatic focus.

The next section will give an overview on the types of authors that published texts on medical topics in the RS journals and on the types of health-related issues

---

3.   The "text type" was extracted from metadata files obtained from the RS. This manually indexed variable has categories such as abstract, article, astronomical observation, biography, book review, catalogue, editorial, experiment, lecture, letter, list, report, speech, etc. Nevertheless journal texts can have elements from several of these options and most texts were assigned broad labels such as "article". During the LModE period there were no clear-cut text-type categories yet, and the structural conventions of journal articles have to be seen in an evolving context. Most LModE medical texts in the RSC fall into the text type categories of "articles", "letters", "book reviews" or "abstracts".

that played a role in these journal papers. Medical research is an important and rather interdisciplinary field in the RSC. LModE medical publications in the RSC were predominantly case reports or reports on small-scale experiments. They will be illustrated with examples of their typical linguistic structures.

3. Late Modern English medical texts in the RSC

3.1 Authors of medical texts

The time span between the end of the seventeenth and the beginning of the twentieth century was characterized by scientific progress in anatomical and physiological research and by developments in medical and surgical practice reported to the RS. Therefore, this section focuses on articles from these fields in the time span covered by the *RSC 6.0 Open*. These articles will also be compared with later, more recent publications included in the *RSC 6.0 Full*. The metadata of the RSC texts provide users with powerful ways to identify particular groups of texts within the data. In various other specialized corpora that have been used for the analysis of medical discourse, finding texts on medicine written by experts in the field does not present the same challenges if such corpora include only texts from medical publications. In less specialized corpora that include some texts which may be relevant for the analysis of medical discourse the respective texts have typically been located through manual work or basic queries within the texts such as the search for individual words related to specific medical fields, treatments, diseases etc.

Metadata on the background of the authors of the RSC texts are one possible source for identifying texts by certain types of authors. Most texts have at least one RS Fellow or Foreign Member as their authors. Detailed information on all Fellows and Foreign Members has been extracted from the publicly available *Royal Society Fellows Directory* in order to facilitate searching for specific biographical information that may be difficult to retrieve systematically from the content of the texts in the RSC alone, e.g., full name, gender, nationality, date of birth and place of residence. Discourse participants with a medical background as potential authors of medical articles in the RSC can often be identified via information contained in categories such as "Activity" and its subcategories "Research Field", "Profession" and "Career" in the *Fellows Directory*. Table 2 shows a sample extract of a record from the *Fellows Directory* with biographical data of one author in the RSC.[4]

---

4. Information on medical qualifications and memberships is often given in an abbreviated form, e.g., MD = Doctor of Medicine, FRCP = Fellow of the Royal College of Physicians, MRCS = Member of the Royal College of Surgeons.

Table 2. Extract from *Fellows Directory* with biographical data for one
of the authors of medical texts in the RSC

Category	Value
Surname	Parkes
Forenames	Edmund Alexander
Dates of Existence	1819–1876
Nationality	British
Address	4 Sussex Place, Hyde Park, London (1861)
Activity	Profession: Physician, Surgeon
	Research Field: Hygiene
	Education: Christ's Hospital; University College Hospital, London. MB (1841) MD (1846, London)
	Career: […] Professor of Clinical Medicine, University College Hospital, London (1849) […]
	Memberships:
	Royal College of Surgeons, Royal College of Physicians
Code	NA6029

Another example of an article with medical implications which can be found in the RSC by the use of such biographical metadata is a text by two Fellows – John T. Cash (Profession: Physician, Research Field: Pharmacology; Career: Professor of *Materia Medica* at the University of Aberdeen) and Wyndham R. Dunstan (Professor of Chemistry, Director of the Research Laboratory of the Pharmaceutical Society) – with the title "The pharmacology of pseudaconitine and japaconitine considered in relation to that of aconitine" (ID: rspl_1901_0063).[5]

The biographical sketches of RS Fellows are a useful resource to locate articles from particular research areas in the RSC, but sometimes additional indicators are needed to identify LModE medical articles in cases where researchers would represent several research fields. Some authors were originally trained as health professionals and carried licenses as physicians or surgeons but mainly worked in different fields and may have published on a variety of topics, e.g., William H. Wollaston, who is best known for his chemical work.

Apart from using metadata in connection with the *Fellows Directory* on the educational background, research field and professional activities of text authors, many medical articles in the RSC can also be retrieved by searching for information on the professional background and the organizational or academic affiliations of the authors in the corpus texts themselves (see Table 3). In cases where some information on the medical background of text authors is given explicitly, this is

---

5.   ID: Text identification code in the RSC.

Table 3. Examples of information on author background derived from the RSC

Text ID in the RSC	Information on professional background and affiliations of authors	Year
106532	Robert Bland, M.D. Physician-Man-Midwife to the Westminster General Dispensary	1781
105475	Daniel Peter Layard, M.D. Physician to her Royal Highness the Princess Dowager of Wales, Member of the Royal College of Physicians in London and Gottingen	1766
112775	E.A. Parkes, M.D., P.R.S., Professor of Hygiene in the Army Medical School, and Count Cyprian Wollowicz, M.D., Assistant Surgeon, Army Medical Staff	1869
rspb_1905_0051	H. Wolferstan Thomas, M.D., C.M. […], Liverpool School of Tropical Medicine […]	1905

indicated directly after the title of the article, and various abbreviations for quali-
fication and memberships are used.

Among the most prolific authors in the RSC with around 30–200 submitted ar-
ticles (cf. Fischer et al. 2018) we find several medical professionals from the LModE
period, e.g., the surgeon Everard Home with the highest number of single-authored
papers, many of which deal with medical topics, and the above-mentioned William
H. Wollaston, a physician, chemist and physicist who published works on a num-
ber of subjects including medical practice. Early authors of medical texts in the
RSC at the beginning of the LModE periods were typically surgeons or physicians,
occasionally also other types of authors with an interest in medical practice and
new information on causes, effects and treatments of diseases. Later, the authors of
medical articles in the LModE period were predominantly specialized academics
affiliated with medical schools or physiology departments of higher education in-
stitutions, sometimes with some practical medical experience as well (e.g., Arthur
Jacob, M.D., Demonstrator of Anatomy and Lecturer on Diseases of the Eye in the
University of Dublin).

Particular queries allow the retrieval of medical texts with a certain number
of authors or by particular types of authors, e.g., with a certain nationality, birth
place or gender. Papers with women as authors or co-authors, for instance, can
be identified via queries for author information within the corpus texts, or in the
metadata where female first names are either spelled out or "Miss" / "Mrs" appears
in front of the initials of the first names of female authors or co-authors. The first
names of male authors or co-authors are typically reduced to initials.

It does not come as a surprise that medical discourse in LModE publications
was essentially a "male discourse". Female authors who have published in RS jour-
nals are generally underrepresented compared to the number of women working

in the sciences, even in the PDE period (Fyfe and Røstvik 2018: 161).[6] One woman, Dorothy Moyle Needham (1896–1987), stands out in the RSC as co-author of various research articles from the 1920s onwards on chemical processes in biological systems. Most of her publications in the RSC were written together with the renowned biochemist Joseph Needham. Other examples of health-related scientific publications with female co-authors are multi-authored papers such as "A discussion on the application of quantitative methods to certain problems in psychology" (rspb_1938_0035, with British industrial psychologists May Smith and Winifred Raphael among the co-authors) and "The production of cancer by pure hydrocarbons" (rspb_1940_0046, with Ernest Kennaway, his wife Nina Kennaway, and Mrs A.M. Robinson among the co-authors).

Searches for articles with a specific number of authors in the RSC confirm that the proportion of multi-authored papers has generally increased over time. Medical articles written by research teams have become a common form in the PDE part of the RSC, where it is not unusual to find research articles related to medical topics with four or more authors (e.g., "Ultrasonics in medical diagnosis", a nine-author paper, rsta_1979_0053). Most medical texts in the LModE part are reports either submitted by individual authors who were Fellows of the RS or by pairs of non-members and Fellows of the RS while the non-members often were the main authors and the latter mainly acted as "communicators".[7] Some prominent

---

6. Medical articles by women are very rare across the whole dataset, but several articles with female authors – especially as co-authors in later papers with higher number of authors – still have some direct or indirect relevance for human and/or animal health, e.g., articles on animal experimentation by Audrey U. Smith from the 1940s onwards. Among the first women admitted as Fellows or Foreign Members of the RS with a background in biosciences and medicine were, for instance, the American geneticist Salome Gluecksohn-Waelsch (1907–2007) and Rosalind Venetia Pitt-Rivers (1907–1990), a British expert on thyroid hormones. However, there is no article by Gluecksohn-Waelsch in the RSC. Pitt-Rivers co-authored a text that can be found in the *RSC 6.0 Full* – an obituary for Charles Robert Harington, a former director of the National Institute for Medical Research with whom she had worked.

7. Until 1990, in order to submit a paper to the Royal Society one had to be a Fellow. Other potential authors needed to find a sponsor or patron among the Fellows who was willing to "communicate" a paper on their behalf. In LModE, the editorial processes were still strongly embedded in the RS community. The Fellows represented the mainstream of British science. When Fellows communicated papers by outsiders, they acted as strong gatekeepers to the editorial process. Therefore, the historical corpus part is particularly rich in reports on the scientific achievements of the network of the London-based RS. There is a certain dominance of British and Commonwealth scientists and a few other outstanding international researchers and the way they used English in academic and medical discourse. After the LModE period, the number of papers submitted by non-members increased considerably. Although the editors continue to be Fellows, the extremely close links between RS journals and fellowship have largely disappeared (cf. Fyfe et al. 2020).

Fellows steered a large number of papers by non-Fellows through the publication process, often without having contributed to the actual research themselves (cf. Harrison 1989: 112).

3.2   RSC texts with medical implications

The most efficient method for identifying articles relevant for medical, health-related and neighbouring disciplines in the LModE part of the RSC is the use of text topic information in the metadata. Information on disciplines, topics or thematic keywords for the texts was not part of the original dataset. Therefore MALLET (McCallum 2002) was used for obtaining a diachronic probabilistic topic model of the LModE part (Fankhauser et al. 2016). On the basis of this model, an approximate classification scheme for scientific topics and academic disciplines represented in the data has been obtained. This makes it possible to effectively identify groups of texts such as articles on topics from the life sciences or more specifically on medical issues.[8] On the basis of the subset of texts gained by this method, the linguistic characteristics and communicative practices of medical research papers and other articles from closely related fields can be identified systematically to gain new insights into their development in LModE.

Topic modelling has also been applied across *RSC 6.0 Full*. Additional challenges were encountered with this method: due to the fact that a high number of texts from the twentieth century from A and B journals are characterized by the increasing specialization and diversification of the mathematical, physical and life sciences, only a low percentage of all documents in the PDE part is related to medical research. These texts can therefore better be identified by use of other methods. Most of these newer texts on medical topics were published in the B journals of *Phil. Trans.* and *Proc.* and report on studies from emerging fields or interdisciplinary research with medical implications. The B journals cover a broad spectrum of life science topics, e.g., health, disease and epidemiology, biochemistry, cellular biology, developmental biology, genetics, immunology, neuroscience, physiology, structural biology and taxonomy. The A journals with a main focus on mathematics and physics occasionally include articles on medical topics, e.g., in the areas of biochemistry, biomedical engineering, medical computing and medical chemistry. In the *Notes and Records* and the *Biographical Memoirs* journals, various PDE texts are dedicated to the history of medical research in the LModE period (e.g., "The

---

8.   Topic modelling is also used by Schneider (this volume), who demonstrates how this method can be applied to a large dataset consisting only of medical texts. This helps to gain an overview of the most prominent subtopics of medical discourse in a diachronic dataset representing a broad range of sources, text types and discourse participants across several centuries.

weather and diseases: some eighteenth-century contributions to observational me-
teorology", rsnr_1952_0017 or "A respectable mad-doctor? Dr Richard Hale, F.R.S.
(1670–1728)", rsnr_1990_0017).

Table 4 shows some of the LModE text topics obtained by topic modelling
with some of their most characteristic words. A total number of 24 topics has been
determined as an appropriate number of topics in the model created for this data-
set. A selection of fewer topics would result in very broad categories, while a more
detailed division would result in over-clustering the LModE data into many small
topics with highly similar features. Labels for these groups of texts were assigned
manually. Most topics can be labelled as belonging to specific research areas such
as "Botany" or "Physiology", while some texts are grouped together by the model
due to the use of similar vocabulary in relation to observational or experimental
studies or the use of many foreign language elements, e.g., a high number of Latin
words in medical texts.[9]

Table 4. Examples of text topics and top words associated with each topic
in LModE texts in the RSC

Topic	Words
Botany	*plant leaves tree seeds species…*
Chemistry I	*water iron solution experiments…*
Observation	*made great found parts part make time small…*
Paleontology	*bone part bones teeth upper lower jaw…*
Physiology I	*blood time animal hours heart found food quantity…*
Physiology II	*fibres nerves muscles vessels side muscular…*
Reproduction	*cells form species membrane part …*

Clustering of topics has been applied in order to identify broader research areas
within the topic hierarchy. These are represented by the branches of the dendrogram
in Figure 1 showing the resulting major clusters and the relationship between the
topics based on their similarity and dissimilarity (cf. also Fischer, Knappen and
Teich 2018).

---

9.   Some early medical texts in the *Phil. Trans.* were published entirely in Latin and have been
excluded from the *RSC 6.0*. Medical texts with English as their main language from the eight-
eenth and nineteenth century are often interwoven with quotations from publications such as the
Miscellanea curiosa and other neo-classical and classical sources. References to medicine in the
classical world are most typically found in RSC texts from the eighteenth century with quotations
attributed to anatomists and physicians such as Herophilus, Erasistratus or Hippocrates (e.g., in
"The ambe of Hippocrates for reducing luxations of the arm with shoulder, rectified […]", 1743,
104195). Various medical, anatomical and biological terms were introduced in LModE as calques
from classical or post-classical Latin or Greek terms (e.g., *tuba Fallopii* → *Fallopian tube*).
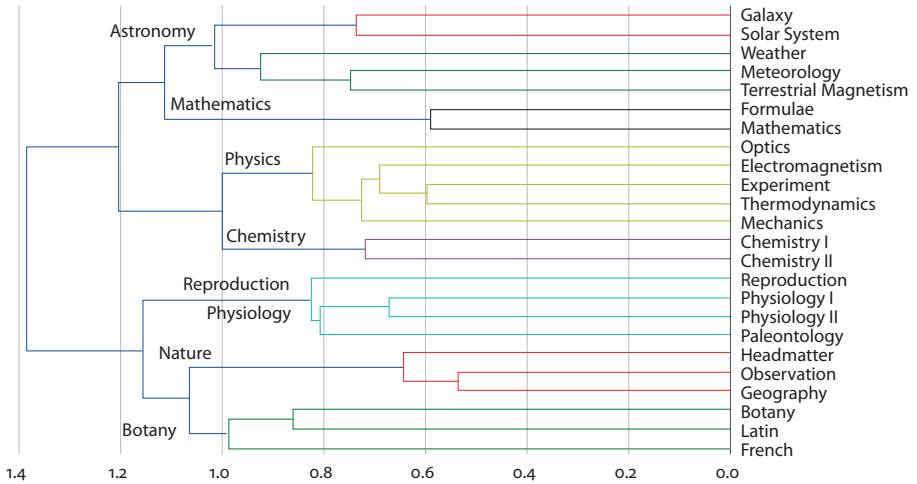
Figure 1. Hierarchical clustering of the topics for LModE texts

The texts have been assigned primary and secondary topics according to their relative importance. RSC texts from the life sciences related to medical discourse in a narrow and larger sense can be identified by taking primary and secondary text topics into account. Manual tests verifying the accuracy of the results of the topic model have shown that medical texts and texts from related fields can typically be found in the groups of texts with Physiology I or II, Reproduction or Observation as their primary and/or secondary topic. They may occasionally also fall into the groups of texts with Botany, Chemistry I or II, Experiment, Latin, Optics or Paleontology as one of their topics. In certain time periods of LModE, additional specific topics such as Thermodynamics or Electromagnetism can be found among the primary and secondary text topics of medical texts; these include articles on experiments under particular temperature conditions or on the therapeutic effects of electricity from the second half of the eighteenth century. Table 5 shows some examples of titles of LModE medical texts identified in the corpus via their assigned topics.[10]

---

10. Regardless of which method is used, there will always remain a certain difficulty to determine clear-cut boundaries for texts with implications for the development of the medical sciences, particularly from a diachronic point of view. Among the articles with topics such as "Physiology I", for instance, are Antoni van Leeuwenhoek's letters on his discovery of bacteria (*animalcule* – "little animals") from the end of the seventeenth century (RSC text ID: 101758, among others). The discoveries by this merchant with an interest in microscopy were doubted by natural scientists for a long time. His findings were confirmed by others later; and only towards the end of the LModE period, bacteria were associated with diseases. Today Van Leeuwenhoek is considered a pioneer of microbiology. Without this contextual knowledge, the corpus metadata on his profession and research field and those gained from topic modelling as well as the qualitative reading of his texts are not sufficient to show why these texts were important for the history of medicine.

Table 5. Examples of medical texts in the *RSC 6.0 Open*

Text ID	Title	Year	Author	Primary and secondary topic
102230	"Lumbricus Teres, or Some Anatomical Observations on the Round Worm Bred in Human Bodies"	1683	Edward Tyson	Observation, Physiology II
103316	"An Account of an Extraordinary Tumour or Wen Lately Cut off the Cheek of a Person in Scotland"	1717	Thomas Bower	Physiology I, Observation
103406	"An account of a fœtus, that continued 46 years in the mother's body"	1721	Johann G. Steigertahl	Physiology I, Observation
103793	"An extraordinary case in Surgery"	1724	Joseph Atkinson	Physiology I, Observation
104068	"A Letter from T. Madden, M.D. of Dublin, to Cromwell Mortimer, M.D.R.S. Secr. Giving an Account of Two Women Being Poisoned by the Simple Distilled Water of Laurel-Leaves, and of Several Experiments upon Dogs; By Which It Appears That This Laurel-Water is One of the Most Dangerous Poisons Hitherto Known"	1731	T. Madden	Physiology I, Observation
104908	"An Account of Those Malignant Fevers That Raged at Rouen at the End of the Year 1753 and the Beginning of 1754"	1755	Claude N. Le Cat	Physiology I, Observation
106277	"Observations on Respiration, and the Use of the Blood"	1776	Joseph Priestley	Chemistry I, Physiology I
112775	"Experiments on the Effect of Alcohol (Ethyl Alcohol) on the Human Body"	1869	Edmund A. Parkes, Cyprian Wollowicz	Physiology I, Chemistry II

The majority of LModE medical texts are individual case reports or reports on small-scale experiments with relevance to human and/or veterinary medicine. The earliest corpus texts are rather short reports in epistolary form. Lehto and Taavitsainen (2019: 89, 109) state that medical case reports have continuity throughout the history of English medical writing from the late medieval period to the present but that their functions and linguistic realizations vary in different periods. In the corpus of *Late Modern English Medical Texts* (LMEMT), experimental medical reports became common during LModE (Lehto and Taavitsainen 2019: 109). In the RSC, experimental reports were common already among earlier medical research articles from the seventeenth century (e.g., "An account of

the experiment of transfusion, practised upon a man in London", 1666, 101241; "Some New Experiments of Injecting Medicated liquors into Veins, together with the considerable Cures perform'd thereby", 1666, 101243). Experimental reports from the RSC throughout the LModE document numerous studies carried out by medical scholars on themselves, e.g., to understand the effects of chemical substances or electricity (e.g., "Extracts of two letters from Dr. John Lining, physician at Charles-Town South Carolina, to James Jurin, M.D.F.R.S. giving an account of statical experiments made several times in a Day upon himself, for one whole Year, accompanied with meteorological observations […]", 1742, 104204; "An extract of a letter from Mr. John Henry Winkler, Graec. & Lat. Litt. Prof. publ. Ordin. at Leipsick, to a friend in London; concerning the effects of electricity upon himself and his wife", 1746, 104806).

In the LMEMT with its different text categories, instruction and giving advice on how patients should be treated was the most common purpose of writing in case reports (Lehto and Taavitsainen 2019: 89), while the authors of medical research articles in the RSC seemed particularly eager to report on unusual cases they had witnessed or on experiments they could present as newsworthy events (see also Taavitsainen's (2015: 141) analysis of samples from the *Phil. Trans.*). Earlier LModE texts in the RSC generally have a stronger focus on medical practices, while later texts more often focus on biological systematics and anatomy.

The results obtained from topic modelling have led to the conclusion that the RSC is rich in case studies documenting surgical, curative and prophylactic practices and treatments as well as medical experiments. Early medical texts from the corpus usually address topics of broader interest that were suitable for RS journal readers with diverse academic interests and professional backgrounds, while later texts represent an increasingly specialized field.

3.3   Linguistic features of medical articles

This section will give an overview on various textual features of medical texts in LModE in the RSC, considering typical morphological, lexical, collocational and phrasal text elements as well as semantic relations. As discussed above, there can probably be no clear cut-off point on which everybody would agree for what constitutes a medical text, as it includes a rich variety of activities, participants and discourse contexts that differed across time. Register conventions for medical research articles written in English have evolved throughout the LModE period. In the RSC, using information from topic modelling on the respective primary and secondary topics of texts is the most efficient method to identify texts of particular scientific fields. In this section, I analyze the linguistic features of text groups whose primary

and/or secondary topics identify them as representing medical-, health-related and neighbouring disciplines, i.e. "Physiology I/II" and "Observation".[11]

One of the textual elements in medical articles that show structural change over time are the titles. The titles of LModE articles may include quite lengthy summaries of the respective research topics, as illustrated in Example (1):

(1) "Some Calculations of the Number of Accidents or Deaths Which Happen in Consequence of Parturition; And of the Proportion of Male to Female Children, as Well as of Twins, Monstrous Productions, and Children That are Dead-Born; Taken from the Midwifery Reports of the Westminster General Dispensary: With an Attempt to Ascertain the Chance of Life at Different Periods, from Infancy to Twenty-Six Years of Age; And Like-Wise the Proportion of Natives to the Rest of the Inhabitants of London"   (1781, 106532)

Such long titles also reveal information on the author and recipient and on the type of the academic paper (e.g., "Letter from T. Madden, M.D. of Dublin, to Cromwell Mortimer, M.D.R.S. Secr. […]", 104068, see Table 5).

The text titles show a lot of structural variation in the beginning of the LModE period. After 1800, more informational and content-oriented, condensed noun phrase structures become the most frequently used patterns in titles, while additional information is more regularly incorporated into other text parts such as abstracts or subheadlines. An important function of titles of medical articles from the LModE period is to attract attention and stimulate curiosity by the use of evaluative adjectives, such as *extraordinary, remarkable* or *dangerous* (see Table 5).

Bizzoni et al. (2019) have shown that diachronic word embeddings are a powerful tool to investigate changing language use in the domain of science. By embedding the words from the RSC into a vector space in a sequence of time slices we can model lexical choice in context as well as lexical-semantic change over time. A publicly available interactive visualization tool provides such co-occurrence information for the RSC, represented on a 2D plane, and can be used to identify clusters of semantically similar lexemes from the medical and related life sciences (cf. Fankhauser and Kupietz, n.d.). Figure 2 illustrates an example of the semantic neighbourhood of the word *anatomist* in LModE texts. Colour denotes frequency change (violet = decreasing, red = increasing), and bubble size encodes proportional relative frequency in a given time interval. Words with similar usage contexts are positioned closely together. Paradigmatic change can be assumed for islands with words of similar colour. The histogram on the right shows frequency changes of individual words in frequency per million tokens (FpM).

---

11. To avoid scarcity of data for particular structures the text groups will be summarized into broader categories, e.g., only primary topics will be taken into account in some cases.

Figure 2. Zoomed-in region of the semantic neighbourhood of *anatomist* in LModE

In the visualization in Figure 2, *anatomist* co-occurs with other lexemes for specialized professionals from related disciplines and other academic fields (e.g., *physiologist, botanist, chemist*)[12] but also with lexemes referring to people performing activities that are relevant for medical research as we have also seen in the results from topic modelling (e.g., *experimenter, observer*). This is in line with the results of Taavitsainen's (2015: 140) keyword analysis, in which she showed that lexical items such as *observation* and *experiment* as indicators of main scientific activities in this field were typical of the *Phil. Trans.* medical articles in comparison to other texts. Other examples of classes of semantically close lexical elements occurring in respective expanding or shrinking word clusters in the visualisation are symptoms, anatomical terms, types of diseases as well as adjectives from medical contexts, e.g., *bony, capsular, cellular, cartilaginous, dermal*, and *vascular*, many of which are built from Latinate morphemes. This is illustrated in Figure 3, where several eponymous adjectives ending in the derivational morphemes *-ian* or *-ean* attached to the names of physicians, physiologist and anatomists can also be found within close vicinity to each other in the group of adjectives from medical contexts (e.g., *Eustachian, Fallopian, Graafian, Haversian, Purkinjean*).

Building on the information gained from word embeddings on the development of groups of lexemes during the LModE period, we can develop CQP-queries focused on items from these clusters to retrieve larger term and phrase patterns in medical discourse. A query on adjectives from medical contexts followed by nouns, for instance, reveals that such structures in the eighteenth century were less frequent and less specific than in the nineteenth century. Among the most frequent collocations with this pattern in the eighteenth century are examples such as *bony part, cellular substance*, and *cartilaginous nature*. In nineteenth century texts we

---

12. Spelling variants and plural forms are represented as separate items in this visualization.

Figure 3. Zoomed-in region of the semantic neighbourhood of the adjective *vascular* in LModE texts in the RSC

find more specific adjective + noun pairs which are used as scientific terms such as *cellular tissue, dermal membrane* or *vascular system*. Eponymous adjective + noun constructions are frequently used in medical terminology and in anatomical and biological nomenclature across LModE medical texts, e.g., *Eustachian tube, Fallopian tube, Graafian vesicle, Haversian canals, Haversian system,* and *Purkinjean vesicle* (cf. Menzel 2021). Many eponyms of this type were coined in LModE and continue to be used in PDE texts in the RSC, e.g., *Wolffian duct* (also known as *mesonephric duct*), named after a physiologist from the eighteenth century.

Other prominent features of LModE medical case reports in the RSC are evaluative adjectives before nouns. Their occurrence is not limited to text titles. This becomes obvious when combining metadata queries targeted to identify medical articles with queries for selected linguistic structures in these groups of texts, e.g., searching for adverb + adjective + noun structures in texts with "Physiology I/II" and "Observation" as primary and secondary topics in LModE texts. The word cloud in Figure 4 captures some of the most frequent collocations occurring in this pattern in LModE texts with these topics. The texts were still characterized by a large proportion of rather general language vocabulary compared to PDE medical texts.

Evaluative adjectives were also frequently used in the pattern *It/it is ADJ that/to* in LModE (see also Hiltunen, this volume, on the role of this pattern in stance marking). Figure 5 shows the diachronic development of the frequency of this pattern in selected groups of texts, e.g., in medical research papers with "Physiology I/II" or "Reproduction" as primary topics.

Figure 4. Most frequent adverb + adjective + noun structures in LModE texts with "Physiology I/II" as primary and "Observation" as secondary topic

Figure 5. FpM of *It/it is ADJ that/to* pattern in first 200 years of the RSC across texts with selected primary topics[13]

A general trend in all disciplines analyzed in Figure 5 is that the *It/it is ADJ to/that* pattern goes up in frequency from the period of 1750–1799 onwards.[14] In relation to

---

13. Frequencies have been normalized in relation to the respective size of each domain-specific subcorpus.

14. Similar structures to this pattern not considered in this analysis can be found with broader types of queries allowing other forms of BE (e.g., past form, subjunctive or *-ing* form), comma before *that* (e.g., *It/it is manifest, that*), adverbs or negation markers before the adjective, etc. Structures that are irrelevant for this evaluative pattern (e.g., not followed by clausal structure, but by noun phrase after *to* as in *it is parallel to the axis*) were excluded via more specific queries.

this aspect, medical discourse does not seem to have a special status in the context of other disciplines in the RSC. The writing style across disciplines seems to become more homogeneous over time after the RS's statutes had granted the right to select papers to be published to a committee in the 1750s.[15] In texts with medical topics this pattern predominantly has an evaluative pragmatic function. The most popular lexical choices for adjectives in this pattern in the texts taken into account in this analysis related to human and veterinary medical research are *difficult, evident, impossible* and *probable*. Some examples are given in Table 6.

Table 6. Examples of *It/it is ADJ to/that* structures in LModE medical texts in the RSC

ID	Example	Text title	Author, year
107507	"[…] it is difficult to separate it to any extent […]"	"An account of a Membrane in the eye, now first described"	Arthur Jacob, 1819
107340	"It is evident that this poison acts in some way or another on the brain […]"	"Experiments and Observations on the Different Modes in Which Death is Produced by Certain Vegetable Poisons"	Benjamin Collins Brodie, 1811
112342	"[…] it is impossible to do more than speculate […]"	"On the Brain of a Bushwoman; And on the Brains of Two Idiots of European Descent"	John Marshall, 1862
109881	"It is probable that the ovum is retained there for several days […]"	"On the Passage of the Ovum from the Ovarium to the Uterus in Women"	Everard Home, 1815

Qualitative observations of the data show that linguistic precision, objective style and the features of plain English have become more dominant textual features over time in all disciplines However, LModE texts that refer to medical concepts seem to be marked by particularly elaborate and subjective formulations and tend to remain more personal in stance than research articles from other disciplines. One of the reasons for this may be a register-related convention to explicitly evaluate information presented in factual medical reports. Another reason might be that many medical professionals mentioned in the RSC often held multiple influential positions in society. The use of positive evaluative adjectives in noun phrases is an important means of giving credit to medical professionals and of expressing appreciation and politeness towards colleagues. When authors mention medical

---

15. Previously this selection was made by a secretary of the Society. The *Phil. Trans.* were licensed by the RS from the beginning, but had started as a quasi-private venture under the responsibility of individuals, e.g., Henry Oldenburg.

professionals in the LModE period, they typically use phrases such as *a(n) eminent / celebrated / ingenious / judicious / able / distinguished / famous / good physician, a very skillful man-midwife, an eminent apothecary and botanist*, and *an incomparable anatomist and learned physician.*

The RSC data confirm observations on various linguistic aspects for the whole LModE period that have been observed on the basis of smaller samples, e.g., a selection of eighteenth century case reports from the LMEMT in Lehto and Taavitsainen's study (2019: 91), in which they observed that the authors' presence is prominently intertwined with the patients' stories and the texts are rich in narrative elements. LModE medical texts in RS publications have a special status as their language and structure were kept simpler than in texts from other disciplines, probably in order to attract readers from other fields as well. They also have different linguistic features than PDE medical research articles, which are often said to be hard to understand and to be accessible only to a narrow audience so that many of the handbooks on medical writing mentioned above emphasize the need for using plain English and to avoid medical jargon and grammatical complexity whenever possible in order to reach the widest reasonable audience and enhance the spread of new medical knowledge. In contrast to other disciplines, first-person narratives of newsworthy cases were still very common among the later medical texts from the nineteenth century. They are characterized by a slightly less involved, more informational and content-oriented style and more nominalizations and extended noun-phrases than earlier scientific texts, which is in line with previous findings from Atkinson (1999: 110ff.), however to a lesser extent than articles from other disciplines. A real shift to non-human technical facts and larger empirical studies takes place only in the PDE medical texts from the corpus. After the LModE period, more forms of dense encoding are used and we see a potentially higher information density of whole texts. Neoclassical terms and morphemes as parts of nouns and adjectives contribute to high information load and highly specialized terminology in modern medical discourse as in *akinetopsia* or in *transcranial*. Neither example is attested in dictionaries such as the Oxford English Dictionary (OED), indicating that specialized corpora like the RSC serve as a valuable resource for specialized lexis (cf. Menzel 2018). The importance of complex term patterns, e.g., consisting of several adjectives and nouns such as *cerebral visual motion blindness*, and of reduced forms, e.g., *TMS* (transcranial magnetic stimulation) or *V1/V5* (brain areas: primary visual cortex / middle temporal visual area) also increases considerably after LModE in the RSC texts.

4.   Summary and conclusion

The analysis of the *RSC 6.0* illustrates that LModE was a productive time span in the medical output of the Royal Society. The corpus is a rich resource for text-oriented register analyses and socio-pragmatic studies on variables that influence linguistic choices in scientific journal articles from medical and related sciences from this time period. As an exhaustive domain-specific corpus that contains nearly all texts from a long time span published in a selection of several scientific journals, the RSC is a unique resource for historical linguists, sociolinguists, historians of science and researchers with an interest in English for specific and academic purposes. As a tool for research and teaching, the RSC provides a possibility to observe the linguistic characteristics of academic journal articles, some of which appear to be highly dependent on the socio-historical context and might have changed considerably over time while others seem to exhibit a certain degree of transhistorical stability.

Medical texts have a special status. In the RSC they reflect research from an interdisciplinary field. Medical discourse in LModE academic journal articles is not limited to academic researchers from the medical sciences but also includes other discourse participants with an interest and experience in health care related issues. Although the author's place generally changes in scientific publications over the LModE period, medical articles remain rather author-centered, strongly narrative, detail- and stance-oriented and less abstract than articles from other disciplines. Medical articles in the RSC frequently deal with topics of broader interest to attract readers with diverse academic and professional backgrounds. Textual reference to other published medical works in LModE frequently means reference to classical sources. The description of events from a quantitative and technological perspective only becomes important in PDE medical texts.

This paper has demonstrated how the RSC may be used to identify and verify certain features of historical medical research articles that have been suggested in the literature on the basis of smaller samples. The analyses demonstrate how the RSC facilitates the systematic identification of medical articles within a large multi-disciplinary corpus and illustrate various aspects of the communicative practices in medical research papers. Over the centuries, many life sciences publications covering a wide variety of medical topics have appeared in RS journals. The results of quantitative and qualitative analyses in the RSC set the development of English medical discourse in a broader historical framework and contribute to a better understanding of the special status of medical texts. These results also have implications for medical discourse analysis with regard to other languages as medical academic discourse and its specialized jargon in Western societies share certain characteristics due to a common linguistic, historical and cultural heritage and a long-standing tradition of international exchanges in the field of medicine.

Acknowledgements

References

Atkinson, Dwight. 1992. "The Evolution of Medical Research Writing from 1735 to 1985: The Case of the *Edinburgh Medical Journal*." *Applied Linguistics* 13 (4): 337–374. https://doi.org/10.1093/applin/13.4.337

Atkinson, Dwight. 1999. *Scientific Discourse in Sociohistorical Context: The Philosophical Transactions of the Royal Society of London, 1675–1975*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Bellés-Fortuño, Begoña. 2016. "Popular Science Articles vs. Scientific Articles: A Tool for Medical Education." In *New Insight into the Analysis of Medical Discourse in Professional, Academic and Popular Settings*, ed. by Pilar Ordoñez-Lopez, and Nuria Edo-Marza, 55–78. Bristol: Multilingual Matters. https://doi.org/10.21832/9781783096268-005

Biber, Douglas, and Edward Finegan. 1994. "Intra-textual Variation within Medical Research Articles." In *Corpus-based Research into Language*, ed. by Nelleke Oostdijk, and Pieter de Haan, 201–222. Amsterdam: Rodopi.

Bizzoni, Yuri, Stefania Degaetano-Ortlieb, Katrin Menzel, Pauline Krielke, and Elke Teich. 2019. "Grammar and Meaning: Analysing the Topology of Diachronic Word Embeddings." In *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change*. Florence: Association for Computational Linguistics. https://www.aclweb.org/anthology/W19-4722/. Accessed 25 November 2019. https://doi.org/10.18653/v1/W19-4722

Crespo-García, Begoña, and Isabel Moskowich. 2015. "A Corpus of History Texts (CHET) as Part of the Coruña Corpus Project." In *Proceedings of Corpus Linguistics 2015*, 14–23. St. Petersburg: St. Petersburg State University.

Degaetano-Ortlieb, Stefania, and Elke Teich. 2016. "Information-based Modeling of Diachronic Linguistic Change: From Typicality to Productivity." In *Proceedings of the 10th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, ed. by Nils Reiter, Beatrice Alex, Kalliopi, and A. Zervanou Berlin: Association for Computational Linguistics. https://www.aclweb.org/anthology/W16-2121, accessed 25 November 2020. https://doi.org/10.18653/v1/W16-2121

Degaetano-Ortlieb, Stefania, and Elke Teich. 2019. "Toward an Optimal Code for Communication: The Case of Scientific English." *Corpus Linguistics and Linguistic Theory* 0(0):1–33. https://doi.org/10.1515/cllt-2018-0088

De Smet, Hendrik, Susanne Flach, Jukka Tyrkkö, and Hans-Jürgen Diller. 2015. *The Corpus of Late Modern English Texts (CLMET), version 3.1: Improved tokenization and linguistic annotation*. Available online at https://hdl.handle.net/21.11119/0000-0002-43F3-0

Fankhauser, Peter, and Marc Kupietz. N.d. *Visual Correlation for Exploring Paradigmatic Language Change*. http://corpora.ids-mannheim.de/diaviz/description.html, accessed 25 November 2019.

Fankhauser, Peter, Jörg Knappen, and Elke Teich. 2016. "Topical Diversification over Time in the Royal Society Corpus." *Digital Humanities 2016 Conference Abstracts*, 496–500. Kraków: Jagiellonian University & Pedagogical University.

Ferguson, Gibson. 2001. "If You Pop Over There: A Corpus-based Study of Conditionals in Medical Discourse." *English for Specific Purposes* 20 (1): 61–82. https://doi.org/10.1016/S0889-4906(99)00027-7

Fischer, Stefan, Jörg Knappen, Katrin Menzel, and Elke Teich. 2020. "The Royal Society Corpus 6.0 – Providing 300+ Years of Scientific Writing for Humanistic Study." *Proceedings of the International Conference on Language Resources and Evaluation 2020 (LREC2020)*. Marseille, France: European Language Resources Association (ELRA). 794–802. http://www.lrec-conf.org/proceedings/lrec2020/index.html

Fischer, Stefan, Peter Fankhauser, and Elke Teich. 2017. "Visualization of Corpus Frequencies at Text Level". *Extended Abstracts of the Corpus Linguistics Conference (CL2017)*. University of Birmingham, UK. 25–28 July 2017. https://www.birmingham.ac.uk/Documents/college-artslaw/corpus/conference-archives/2017/general/paper112.pdf

Fischer, Stefan, Knappen, Jörg, and Elke Teich. 2018. "Using Topic Modelling to Explore Authors' Research Fields in a Corpus of Historical Scientific English." *Abstracts of DH 2018*, 581–584. Mexico City, Mexico. https://dh2018.adho.org/using-topic-modelling-to-explore-authors-research-fields-in-a-corpus-of-historical-scientific-english/

Fløttum, Kjersti. 2006. "Medical Research Articles in the Comparative Perspectives of Discipline and Language." In *Advances in Medical Discourse Analysis: Oral and Written Contexts*, ed. by Maurizio Gotti, and Françoise Salager-Meyer, 251–269. Bern: Peter Lang.

Fyfe, Aileen, and Camilla Mørk Røstvik. 2018. "How Female Fellows Fared at the Royal Society." *Nature* 555 (7695): 159–161. https://doi.org/10.1038/d41586-018-02746-z

Fyfe, Aileen, Flaminio Squazzoni, Didier Torny, and Pierpaolo Dondio. 2020. "Managing the Growth of Peer Review at the Royal Society Journals, 1865–1965." *Science, Technology, & Human Values* 45 (3): 405–429. https://doi.org/10.1177/0162243919862868

Goodman, Neville W., and Martin B. Edwards. 2014. *Medical Writing: A Prescription for Clarity*. 4th edition. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107279179

Halliday, Michael Alexander Kirkwood. 1987. "Language and the Order of Nature." In *The Linguistics of Writing*, ed. by Nigel Fabb, and Alan Durant, 135–154. Manchester: Manchester University Press.

Hardie, Andrew. 2012. "CQPweb – Combining Power, Flexibility and Usability in a Corpus Analysis Tool." *International Journal of Corpus Linguistics* 17 (3): 380–409. https://doi.org/10.1075/ijcl.17.3.04har

Harrison, Andrew John. 1989. Scientific Naturalists and the Government of the Royal Society 1850–1900. PhD dissertation. The Open University.

Jenicek, Milos. 2014. *Writing, Reading, and Understanding in Modern Health Sciences: Medical Articles and other Forms of Communication*. Boca Raton: Taylor & Francis.

Kermes, Hannah, Stefania Degaetano-Ortlieb, Ashraf Khamis, Jörg Knappen, and Elke Teich. 2016. "The Royal Society Corpus: From Uncharted Data to Corpus." *Proceedings of the International Conference on Language Resources and Evaluation 2016 (LREC2016)*, ed. by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, 1928–1931. Portoroz, Slovenia: European Language Resources Association.

Knappen, Jörg, Stefan Fischer, Hannah Kermes, Elke Teich, and Peter Fankhauser. 2017. "The Making of the Royal Society Corpus." In *Proceedings of the 21st NoDaLiDa Workshop on Processing Historical Language*, May 2017, ed. by Jörg Tiedemann and Nina Tahmasebi, 7–11. Gothenburg, Sweden: Association for Computational Linguistics.

Laso, Natalia Judith. 2009. A Corpus-Based Study of the Phraseological Behaviour of Abstract Nouns in Medical English: A Needs Analysis of a Spanish Medical Community. PhD dissertation. University of Barcelona.

Lehto, Anu, and Irma Taavitsainen. 2019. "Medical Case Reports in Late Modern English." In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen, 89–111. Amsterdam: John Benjamins. https://doi.org/10.1075/z.221.05leh

Linares, Oscar, David Daly, and Gertrude Daly. 2017. *Plain English for Doctors and Other Medical Scientists*. New York: Oxford University Press.

McCallum, Andrew Kachites. 2002. MALLET. http://mallet.cs.umass.edu/

MEMT – Middle English Medical Texts corpus, CD-ROM. 2005. *Compilers: Irma Taavitsainen, Päivi Pahta and Martti Mäkinen*. Amsterdam: John Benjamins.

Menzel, Katrin. 2021. "Scientific Eponyms throughout the History of English Scholarly Journal Articles." In *Broadening Perspectives in the History of Dictionaries and Word Studies (Selected Papers from ICHLL 10)*. ed. by Hans Van de Velde and Fredric Dolezal. Newcastle upon Tyne: Cambridge Scholars Publishing, 159–193.

Menzel, Katrin, and Stefania Degaetano-Ortlieb. 2017. "The Diachronic Development of Combining Forms in Scientific Writing." *Lege Artis. Language Yesterday, Today, Tomorrow*, 2 (2): 185–249. https://doi.org/10.1515/lart-2017-0016

Menzel, Katrin. 2018. "Using Diachronic Corpora of Scientific Journal Articles for Complementing English Corpus-based Dictionaries and Lexicographical Resources for Specialized Languages." *Proceedings of, the XVIII EURALEX International Congress: Lexicography in Global Contexts, July 2018, Ljubljana, Slovenia*, ed. by Jaka Čibej, Vojko Gorjanc, Iztok Kosem, and Simon Krek. 363–372. Ljubljana: Ljubljana University Press.

Menzel, Katrin, Jörg Knappen, and Elke Teich. 2021. "Generating Linguistically Relevant Metadata for the Royal Society Corpus." *Research in Corpus Linguistics* 9 (1), 1–18. *Oxford English Dictionary*. http://www.oed.com/. https://doi.org/10.32714/ricl.09.01.02

Pahta, Päivi, and Irma Taavitsainen. 2004. "Vernacularisation of Scientific and Medical Writing in its Sociohistorical Context." In *Medical and Scientific Writing in English 1375–1500*, ed. by Irma Taavitsainen, and Päivi Pahta, 1–22. Cambridge: Cambridge University Press.

Panocová, Renáta. 2015. "Transparency and Use of Neoclassical Word Formation in Medical English." In *Word Formation and Transparency in Medical English*, ed. by Pius ten Hacken, and Renáta Panocová, 73–96. Newcastle upon Tyne: Cambridge Scholars Publishing.

Panocová, Renáta. 2017a. *The Vocabulary of Medical English: A Corpus-based Study*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Panocová, Renáta. 2017b. "Synonymy in Medical Terminology – A Challenge in Translation." In *Translation Studies and Translation Practice: Proceedings of the 2nd International TRANS-LATA Conference*, 2014, Part 1, ed. by Lew N. Zybatow, Any Stauder, and Michael Ustaszewski, 207–212. Frankfurt am Main: Peter Lang.

Roberts, Celia. 2006. "Continuities and Discontinuities in Doctor-Patient Consultations in a Multilingual Society." In *Advances in Medical Discourse Analysis: Oral and Written Contexts*, ed. by Maurizio Gotti, and Françoise Salager-Meyer, 177–195. Bern: Peter Lang.

Royal Society Fellows Directory. N.d. https://royalsociety.org/fellows/fellows-directory/

Salager-Meyer, Françoise. 1999. "Contentiousness in Written Medical English Discourse: A Diachronic Study (1810–1995)." *Text* 19 (3): 371–398.

Stuart, Mark C. (ed.). 2007. *The Complete Guide to Medical Writing*. London: Pharmaceutical Press.

Taavitsainen, Irma, and Päivi Pahta. 2000. "Conventions of Professional Writing: The Modern Medical Case Report in a Historical Perspective." *Journal of English Linguistics* 28: 60–76. https://doi.org/10.1177/00754240022004875

Taavitsainen, Irma. 2011. "Medical Case Reports and Scientific Thought-styles." In *Revista de Lenguas para Fines Específicos 17. Special Issue: Diachronic English for Specific Purposes*, ed. by Francisco Alonso Almeida, 75–98. Spain: Universidad de las Palmas de Gran Canaria.

Taavitsainen, Irma. 2015. "Medical News in England 1665–1800 in Journals for Professional and Lay Audiences". In *Changing Genre Conventions in Historical English News Discourse*, ed. by Birte Bös, and Lucia Kornexl, 135–160. Amsterdam: John Benjamins. https://doi.org/10.1075/ahs.5.06taa

Taavitsainen, Irma, and Turo Hiltunen. 2019. "Towards New Knowledge: The Corpus of Late Modern English Medical Texts". In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen, 1–16. Amsterdam: John Benjamins. https://doi.org/10.1075/z.221.01hil

Taylor, Robert B. 2018. *Medical Writing: A Guide for Clinicians, Educators, and Researchers*. 3rd edition. New York: Springer. https://doi.org/10.1007/978-3-319-70126-4

Teufel, Simone, and Noemie Elhadad. 2002. "Collection and Linguistic Processing of a Large-scale Corpus of Medical Articles." In: *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC2002*, ed. by Manuel González Rodríguez, and Carmen Paz Suarez Araujo 1214–1219. Las Palmas: European Language Resources Association (ELRA).

Survival or death

MINE/MY and THINE/THY variation in Early Modern English medical writing

Terry Walker and Merja Kytö

Mid-Sweden University / Uppsala University

This quantitative and qualitative study examines the chronological development of the possessive determiners MINE, MY, THINE, and THY in the *Early Modern English Medical Texts* corpus (EMEMT), focusing on the influence of phonological environment and text category, as well as noting lexical items with which MINE and THINE commonly occur. The pattern of decline for MINE and THINE, the phonological context in which they are found, and the collocations identified resemble those found in previous research, but rather than occurring in more formal texts as suggested by previous research, these forms predominantly occur in the more "popular" medical text categories, those with a wider readership, and particularly in texts that are more speech-like in nature (cf. Kytö and Walker 2020).

Keywords: possessive determiners, MINE, MY, THINE, THY, *Early Modern English Medical Texts* (EMEMT), chronological development, phonological environment, collocations, formality, "popular" texts

1.   Introduction

The decline of the possessive determiner forms MINE and THINE, ousted by the forms MY and THY, is one of the well-known developments shaping the English language from the Middle English period and through the Early Modern English period. The origin of the variation has been considered basically phonological, the choice of the forms being affected by the initial sound of the word following the *n* sound. The deletion of *n* before consonants was more or less completed by the end of the Middle English period but some of the variation lingered on over the sixteenth century and even beyond (Schendl 1997: 180, with reference to Jespersen 1927: 16.211). Language historians point out that MINE and THINE were kept alive

longest in contexts where they preceded a vowel or *h* but fighting a losing battle against the MY and THY forms that was essentially over towards the end of the Early Modern English period (Barber 1997: 152; Schendl 1997: 180).

Several corpus-based studies have been carried out to depict and account for the process of change from a number of language-theoretical and empirical perspectives (for a survey of a number of representative studies, see Section 2). The present study exploits a valuable data resource hitherto unexplored as regards the variation and development in the use of the possessive determiner forms MY/MINE and THY/THINE. This resource comprises early medical texts covering a period of two hundred years, between 1500 and 1700. The multifaceted nature of early medical texts and the varying readership that these texts were intended for make the material of great relevance for a study of the possessive determiner forms. The tension between the early traditions in medical practice and the new scientific thought styles of the early modern period (Taavitsainen and Pahta 2010) can be expected to be reflected in linguistic usage and the rivalry of old and new forms.

In terms of methodology, our study of MINE/MY and THINE/THY usage integrates elements from historical variationist corpus linguistics and historical pragmatics. We are interested both in the process by which MINE and THINE decline towards the end of the eighteenth century and in the communicative functions of early modern pronoun usage. Therefore, it is important to look into the rivalry of MINE/MY and THINE/THY forms from the variationist frequency perspective but also to invest in qualitative close readings of actual corpus examples and examine usage in the context of, for example, genre and style (see Jucker and Taavitsainen 2014: 8–9). "Corpus pragmatics" has been defined as an approach which "integrates the foundational methodologies typifying the fields of corpus linguistics and pragmatics" (Rühlemann and Aijmer 2015: 12), a definition suiting our research design very well. Further, considering the historical perspective, our study falls within what has been referred to as "diachronic corpus pragmatics"; here the qualification "diachronic" refers to the diachronic developments rather than the mere historical nature of the data (Jucker and Taavitsainen 2014: 4). Within this framework, as in historical pragmatics overall, corpora have been used from the start "with the insight that frequencies indicate tendencies and point out useful source materials, but qualitative assessments are also needed" (Jucker and Taavitsainen 2014: 7). When collecting data from EMEMT, we apply the form-to-function mapping principle, and expect to throw light on the developments in the pragmatics of pronoun usage through the lens of medical writing.

2.   Research context

A wide range of general or specialized multi- or single-genre corpora have been used for investigations of MINE/MY and THINE/THY variation. Among such studies are Schendl (1997) on data drawn from the *Helsinki Corpus of English Texts* and a number of self-compiled data collections, as well as Nevalainen and Raumolin-Brunberg (2003) and Raumolin-Brunberg and Nevalainen (2007) on data from the *Corpus of Early English Correspondence* (CEEC). Hilpert (2013) examines a variety of historical language corpora extending from Middle to Late Modern English, and Kytö and Walker (2020) looks at data drawn from the *Corpus of English Dialogues 1560–1760* (CED). Further, corpus studies that pay attention solely to second person singular determiners include Busse (2002), who uses the *Shakespeare Corpus* and a corpus of contemporaneous playwrights' works, and Walker (2007), who exploits three text types in the CED plus supplementary materials.

The possibility of "free variation" between the use of the N-forms and the incoming MY/THY forms preceding vowels by 1600, which was advocated by Barber (1976: 208, cited in Schendl 1997), was questioned and eventually refuted by Schendl (1997). Schendl found that the level of formality and text type influenced the process of change, especially between 1550 and 1600 when the occurrence of MY/THY forms increased significantly in informal texts and when these forms apparently became stylistically marked as informal (Schendl 1997: 188–189). At the same time, the use of the incoming and receding forms before vowels was increasingly influenced by suprasegmental and lexical factors, gradually leaving the use of the N-forms associated with a restricted number of lexical items (Schendl 1997: 188–189).

While in his study Schendl (1997) considered the influence of the extralinguistic factors of time, text type and level of formality, Nevalainen and Raumolin-Brunberg (2003) and Raumolin-Brunberg and Nevalainen (2007) consider one text type, correspondence, and set out to demonstrate the influence of the sociolinguistic factors of gender, rank, and the writers' regional origin on the process of change. They also consider the register, that is, whether the letters were addressed to family or non-family members. In Nevalainen and Raumolin-Brunberg (2003) the focus is on the use of the variant forms in four environments over a period extending from 1410 to 1681: the occurrence of the possessive determiners preceding consonants, vowels, initial *h*, and the frequent lexeme OWN. They find that words with an initial consonant occurred primarily with the variants MY and THY by the end of the 1400s, whereas words with an initial vowel reached the same level of usage only towards the period between 1660 and 1681 (Nevalainen and Raumolin-Brunberg 2003: 62). The lexeme OWN lagged behind the vowel curve and had not caught up by the end

of the period investigated, while the MY and THY forms preceding *h* peaked during the period between 1540 and 1579 (Nevalainen and Raumolin-Brunberg 2003: 62). Due to the predominance in the *Early Modern English Medical Texts* (EMEMT) corpus of male authors representing the professions, we do not consider socio-linguistic factors in the current study. However, it is interesting to note that these factors are indeed salient in the correspondence corpus. Women (and especially upper-ranking women) led the diffusion of the incoming MY and THY variants as of the latter part of the fifteenth century onward, while men continued to support the declining N-forms. However, the differences in the rates of use remained relatively small and may have reflected the influence exerted by the lower ranks, mostly represented by men in the corpus data (Nevalainen and Raumolin-Brunberg 2003: 119–120). Among men, it was the group of social aspirers that maintained the receding N-forms in the period between 1540 and 1579 when the lower ranks took the lead in the use of the variants MY and THY; social aspirers joined the users of these forms in the ensuing period between 1580 and 1619, leaving the upper ranks as the group making most use of the receding N-forms (Nevalainen and Raumolin-Brunberg 2003: 119–120, 142).

Regarding the geographical region factor, it seems that writers in the south continued to use the N-variants to a greater extent than those in the north, the change to the use of the MY and THY forms proceeding from the north to the south. In the south, the Court continued to use the N-forms in the period between 1540 and 1579 when the use of the variants MY and THY was already more or less a given in London (Nevalainen and Raumolin-Brunberg 2003: 180). Geographical region also proved to be the most significant of the factors included in the multivariate analyses carried out on the data for the period between 1460 and 1659. While region was the top-ranked variable in MY/MINE and THY/THINE variation across the period 1460–1639, the register factor (i.e., to whom a letter was addressed) was less significant and was superseded by the gender factor when the change had reached its near completion (with 66% to 85% of the instances represented by MY and THY) (Nevalainen and Raumolin-Brunberg 2003: 198–200).

In terms of coverage and methodology, Hilpert (2013) makes an important contribution to research on variation and change as regards first and second person singular possessive determiners. Using the *Penn Parsed Corpora of Middle English* (PPCME) and *Early Modern English* (PPCEME) as well as the part-of-speech tagging, Hilpert subjected a raw dataset covering 38 different spelling variants to manual analysis. After applying screening criteria, he included 18,882 relevant instances in the analyses, distributed across four periods of Middle English and three periods of Early Modern English (Hilpert 2013: 83). The use of the N-forms declined across the period: the MINE and THINE forms represented more than half of the instances in the earlier periods, and for the last period no more than 0.3 per cent was attested

for MINE and 4.5 per cent for THINE. The N-forms proved much less frequent (2,464 instances, or 13.3 per cent of all examples) than the MY and THY forms (Hilpert 2013: 87). The huge differences in the group frequencies demanded the use of advanced statistical methodology. A VNC (Variability-based Neighbour Clustering) analysis was used to merge neighbouring data points (starting with the most similar) and to provide a data-driven periodization scheme. The linguistic and extra-linguistic factors investigated in the study included, among others, time period, the phonological context to the right of the determiner, stress pattern, formality, style, collocational context, speaker idiosyncrasies, and gender. No annotation on gender or variables such as social class or geographical origin were included in the *Penn Parsed Corpora*; however, regarding these variables, it was possible to manually code the data for the gender variable and include that variable in the analysis. A factor that had not been included in previous studies was the priming effects, that is, whether the author had used a form before (Hilpert 2013: 91). Attention was also paid to whether first and second person forms displayed similarities or differences in their trajectories of change (Hilpert 2013: 94–97). The fundamental difference in the usage of first and second person forms was the rise of the second person form YOUR, which was introduced in singular uses in place of THY/THINE. However, only a minor difference could be attested between the use of the two N-forms: MINE had "a slightly stronger preference for the prevocalic environment" than THINE (Hilpert 2013: 95). The effects of the various factors and their interaction were modelled to display the probabilities with which each example could be identified as an N-form or a MY/THY variant in view of the VNC-generated periodization scheme (Hilpert 2013: 97–102). A complex picture emerged from these analyses, highlighting the need to consider the impact of factors from a variety of angles.

Targeting speech-related texts included in the *Corpus of English Dialogues 1560–1760* (CED), Kytö and Walker (2020) focus on the declining MINE and THINE forms and their fall into near-total obsolescence. The study addresses the chronological stages of development, the text types where the forms survived longest, and the speaker groups that were the last to use the forms. The searches yielded 12,850 examples of MINE, MY, THINE and THY as determiners in the CED as a whole, MINE comprising only 1.5 per cent and THINE only 3.3 per cent of the data; the N-forms fell into disuse from the 1680s onward. Constructed dialogue found in Comedy Drama, Fiction and Handbooks proved to be the category favouring and preserving the N-forms; the text types reporting real speech events (Trials and Depositions) generally promoted the incoming MY and THY variants. As in previous studies, certain lexemes proved to co-occur with MINE and THINE. Among these collocations were words beginning with *h* that are of French origin (*heir, honesty, honour, host*) and particular words beginning with vowels (*own, eye*s, *intent, opinion*); *own* as in, for example, *mine own eyes* was particularly common with the N-forms. The study

also included qualitative analyses focusing on the speaker groups that were the last to use the forms and the contexts in which these occurred.

Summing up, the factors considered to have played a role in the use and development of the first and second person possessive determiners include a wide range of linguistic and extralinguistic factors. Among the linguistic factors are phonological, morphological, stylistic, and collocational factors, while extralinguistic factors include, for example, time period, text type, gender and the social background of the language users, as well as the relationship of a text to different levels of formality and to everyday spoken language. The results of our study of the CED texts (Kytö and Walker 2020) prompted us to consider a follow-up study of medical texts that represent a number of text types with affinities to speech-related and colloquial contexts while they also represent texts originating from the opposite end of the oral-written and formal-informal continuums.

3.   Material and method

The material for the present study was drawn from the electronic text collection *Early Modern English Medical Texts* (EMEMT), a two-million-word register-specific corpus comprising a wide and representative range of medical texts sampled to represent six text categories, and totalling 230 texts (Taavitsainen and Tyrkkö 2010: 57). In terms of chronology, the corpus texts cover the period between 1500 and 1700, with the earliest text dating from 1506 and the latest from 1700 (Taavitsainen and Pahta 2010, Corpus Bibliography: 291ff.). The text categories represented include GENERAL TREATISES AND TEXTBOOKS, TREATISES ON SPECIFIC TOPICS (DISEASES, METHODS OF DIAGNOSIS OR TREATMENT, SPECIFIC THERAPEUTIC SUBSTANCES, MIDWIFERY AND CHILDREN'S DISEASES, and PLAGUE), RECIPE COLLECTIONS AND *MATERIA MEDICA*, REGIMENS AND HEALTH GUIDES, SURGICAL AND ANATOMICAL TREATISES, and the *Philosophical Transactions*. The field of medical writing has "fuzzy edges", so including texts in the corpus and classifying them into categories required the corpus compilers to pay attention to a number of factors. These include chronological coverage, variation in discourse forms (monologues, dialogic texts), vernacularization and multilingualism, target audiences (from educated physicians to lay readership), and continuity from medieval traditions to the early modern period (Taavitsainen and Tyrkkö 2010: 8).

EMEMT is well suited for the purposes of the present study as the general frame of the approach adopted was variationist, "aiming to explore and identify linguistic and language-external factors that condition and determine the choice of a specific variant among the possible alternative ways of saying the same thing"

(Pahta and Taavitsainen 2010: 2). The corpus is part of the *Scientific Thought-styles* project, which pays attention to the long diachrony and the multifaceted sociohistorical setting of the texts. These are also features tallying well with the research design of the present study, the variation and development in the use of the possessive determiners MINE, MY, THINE and THY. Sociohistorical and genre-specific features of the EMEMT texts yield important information for the study of variation in the use of these forms across the Early Modern English period and are useful for making comparisons with results obtained in previous studies on other data sources. Among further features that can be expected to prove of significance to the development of the forms are the relationship of the EMEMT texts to everyday spoken language and the level of formality. While most medical texts in the sixteenth century seem to have been intended for the social elite, texts such as cheap remedy books printed in greater quantities could be thought to have reached the popular end of the readership continuum (Marttila 2010: 107). Furthermore, considering the representation of both monologues and dialogue texts in the EMEMT corpus, the material can be expected to yield sufficient data even though second person uses may be better represented in, for instance, a resource such as the *Corpus of English Dialogues 1560–1760* (CED), dedicated to dialogue texts that yield ample stretches of interactive direct speech.

Searches were carried out on the plain text versions of the corpus files targeting all variant forms of MINE, MY, THINE and THY. All examples of these forms in determiner function were extracted and coded for period (at 50-year intervals), text category, and text. We also coded the ensuing character following the form (a vowel, an *h*, or a consonant) and highlighted the full word following the determiner. Coding for the character rather than the sound was necessary as how certain words starting with *h* were pronounced in the Early Modern English period cannot be judged with certainty (cf. Nevalainen and Raumolin-Brunberg 2003: 61). Our quantitative results are presented in Section 4, followed by a microanalysis of the data in Section 5.

4. Quantitative results

There are 2,237 tokens of MINE, MY, THINE and THY as determiners in EMEMT as a whole (1500–1700), with a clear preference for the MY/THY forms, as shown in Table 1, in line with previous research (see Section 2). The distribution of the data shown in Table 1 is statistically significant at the level $p < .01$ ($x^2 = 15.0463$; $df = 1$; p-value is .000105).

Table 1. The determiners MINE, MY, THINE and THY in EMEMT (1500–1700)

MINE	67 (4.2%)	MY	1,520 (95.8%)
THINE	54 (8.3%)	THY	596 (91.7%)

Table 2 presents the results for the data divided into four 50-year periods. Figure 1 collapses the results per period into N-variants versus non-N-variants. The distribution of the data shown in Figure 1 is statistically significant at the level $p < .01$ ($\chi^2 = 78.1971$; $df = 3$; p-value is < 0.00001).

Table 2. The determiners MINE, MY, THINE and THY in EMEMT per period

	1500–1549	1550–1599	1600–1649	1650–1700
MINE	2 (3.5%)	27 (9.3%)	34 (8.8%)	4 (0.5%)
MY	55 (96.5%)	263 (90.7%)	354 (91.2%)	848 (99.5%)
THINE	36 (13%)	8 (5.4%)	6 (6.7%)	4 (2.9%)
THY	240 (87%)	141 (94.6%)	83 (93.3%)	132 (97.1%)
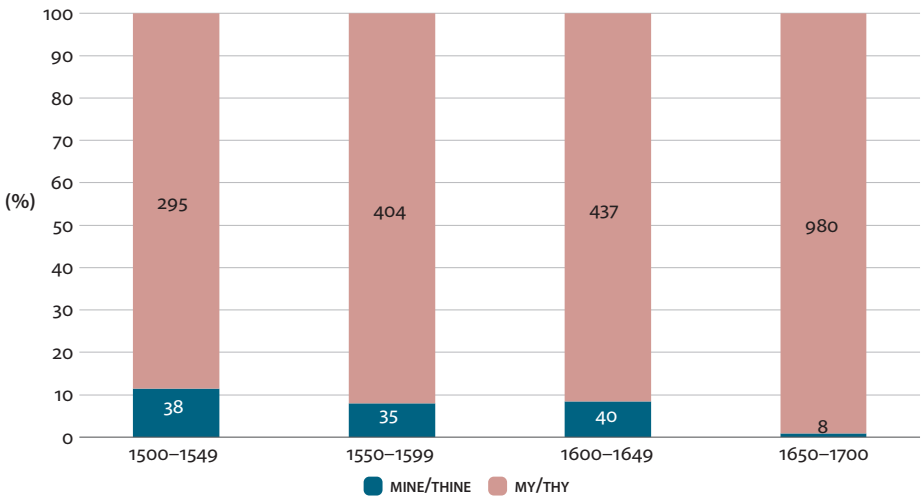
Figure 1. The determiners MY/THY versus MINE/THINE in EMEMT per period

Table 2 and Figure 1 demonstrate that MINE/THINE as determiners, although far less common than MY/THY in the data, remain fairly steady throughout the sixteenth century and the first half of the seventeenth century and do not decline drastically until the second half of the seventeenth century. The percentages, calculated from Figure 1, are 11.4% in the first period, 8.0% and 8.4% in the second and third periods respectively, but a mere 0.8% in the fourth period.

Turning to the phonological factor, we get the results as shown in Table 3.

Table 3. The determiners MINE, MY, THINE and THY preceding a consonant, *h*, or a vowel in EMEMT (1500–1700)

	Consonant	*h*	Vowel	Total (100%)
MINE	0	1 (1.5%)	66 (98.5%)	67 (100%)
MY	1,172 (77.1%)	74 (4.9%)	274 (18%)	1,520 (100%)
THINE	1 (1.9%)	7 (13%)	46 (85.2%)	54 (100%)
THY	462 (77.5%)	73 (12.2%)	61 (10.2%)	596 (100%)
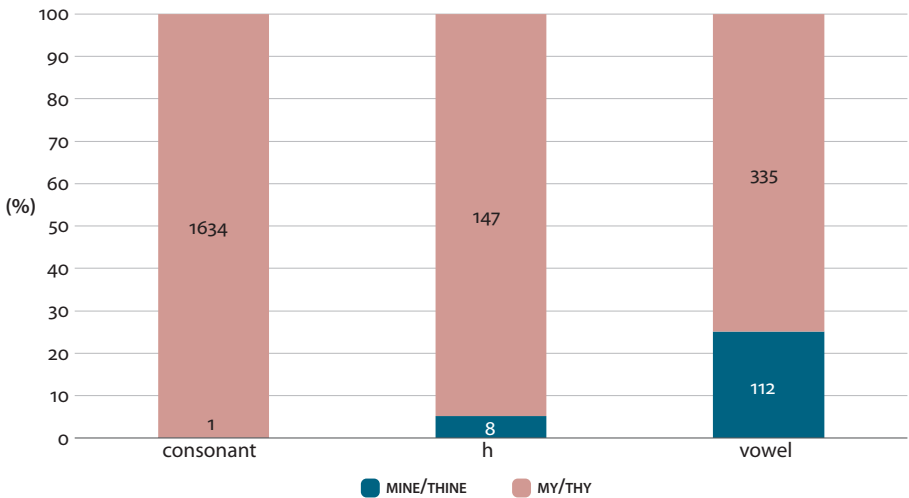
Figure 2. The determiners MY/THY versus MINE/THINE preceding a consonant, *h*, or a vowel in EMEMT (1500–1700)

As can be seen from Table 3 and especially from Figure 2, MINE/THINE preceding a vowel accounts for the great majority, 112 (92.6%), of the 121 examples of MINE/THINE. The distribution of the data shown in Figure 2 is statistically significant at the level $p < .01$ ($\chi^2 = 428.6412$; $df = 2$; *p*-value is < 0.00001). The chronological development is illustrated in Figure 3. The results for MY/THY versus MINE/THINE before a consonant (see Figure 2) tally with the findings of Nevalainen and Raumolin-Brunberg (2003: 62): MY/THY have clearly all but ousted MINE/THINE in this context by the turn of the sixteenth century, with only one example in our data of an N-form preceding a consonant, in a text dating from 1506 (which was first published c. 1490).

As there is only one example of MINE/THINE occurring before a consonant, Figure 3 shows the results in each of the periods for the determiners preceding *h* or a vowel only. The distribution of MINE/ THINE versus MY/THY in EMEMT before a vowel in the four periods shown in Figure 3 is statistically significant at the level $p < .01$ ($\chi^2 = 128.7203$; $df = 3$; p-value is < 0.00001). By contrast, the distribution of MINE/THINE versus MY/THY before *h* in the first three periods (there are no N-forms in the final period) is not significant ($\chi^2 = 4.9356$; $df = 2$; p-value is .084771; moreover, the expected values for the N-forms before *h* are less than 5, so the chi-square result may not be valid). If we recalculate the data from Table 3 to column percentages, the results contrast to some extent with the finding of Hilpert (2013: 95) that "MINE has a slightly stronger preference for the prevocalic environment than second person n-variants" but, as Hilpert specifies, "the difference is not very pronounced". It is THINE in our data that occurs more than MINE before a vowel and before *h*: for THINE (compared with the figures for THY), the percentages are 42.9% and 8.75% for vocalic and *h* environments respectively, whereas for MINE (compared with the figures for MY), the percentages are 19.4% and 1.33%.
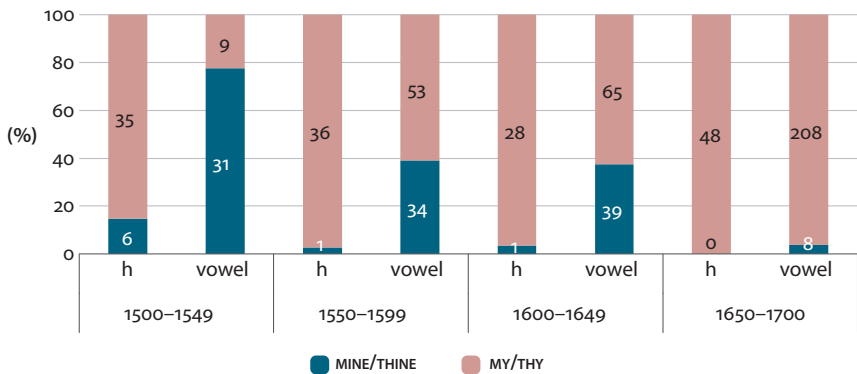
Figure 3. The determiners MY/THY versus MINE/THINE preceding *h* or a vowel in the periods 1500–1549, 1550–1599, 1600–1649 and 1650–1700

Not dissimilar to the findings of Nevalainen and Raumolin-Brunberg (2003: 62) for their correspondence data, MINE/THINE preceding *h* is already in decline before the sixteenth century, as it is infrequent (14.6%, calculated from Figure 3) in the early sixteenth century and rare after 1550 (2.7% and 3.4% in the second and third periods respectively), and not found at all after 1650. By contrast, before 1550, MINE/THINE preceding a vowel is much more frequent than MY/THY (77.5%); the frequencies of MINE/THINE and MY/THY are reversed after 1550, yet the frequency of MINE/THINE remains fairly steady for a century (39.1% and 37.5%) and the N-forms persist (3.7%) up to the end of the seventeenth century, which is similar also to the pattern found by Nevalainen and Raumolin-Brunberg (2003: 62).

The distribution of the determiners according to text category is shown in Table 4 and Figure 4. The distribution of MINE/THINE versus MY/THY in EMEMT per text category shown in Figure 4 (excluding the *Philosophical Transactions* in which there are no N-forms) is not statistically significant at the level $p < .01$, but is significant at the $p < .05$ level ($\chi^2 = 13.0954$; $df = 4$; p-value is .010819).

Table 4. The determiners MINE/MY and THINE/THY per text category in EMEMT (1500–1700)

Category	MINE	MY	THINE	THY
GENERAL TREATISES AND TEXTBOOKS	0	81 (100%)	2 (15.4%)	11 (84.6%)
TREATISES ON SPECIFIC TOPICS	50 (6.4%)	729 (93.6%)	7 (5.6%)	117 (94.4%)
RECIPE COLLECTIONS AND *MATERIA MEDICA*	2 (1.7%)	119 (98.3%)	23 (12.2%)	166 (87.8%)
REGIMENS AND HEALTH GUIDES	6 (5.5%)	103 (94.5%)	20 (8.8%)	206 (91.2%)
SURGICAL AND ANATOMICAL TREATISES	9 (3.2%)	268 (96.8%)	2 (2%)	96 (98%)
Philosophical Transactions	0	220 (100%)	0	0
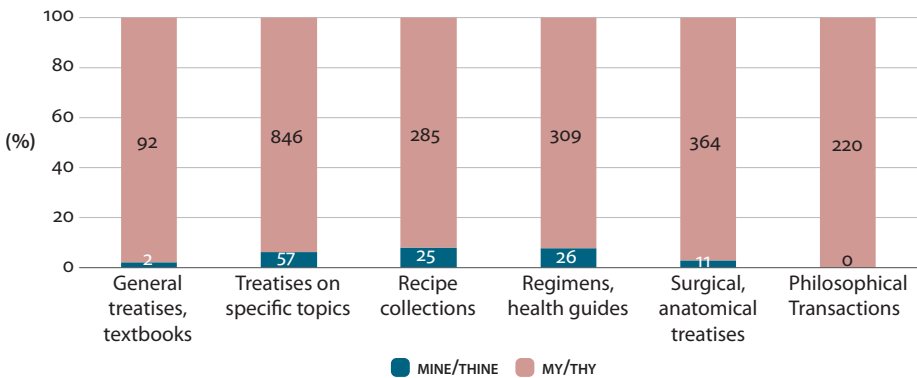
Figure 4. The determiners MY/THY versus MINE/THINE per text category in EMEMT (1500–1700)

Table 4 and Figure 4 reveal that three text categories have a higher ratio of the N-forms than the other three categories. TREATISES ON SPECIFIC TOPICS (6.3%, calculated from Figure 4), RECIPE COLLECTIONS AND *MATERIA MEDICA* (8.1%), and REGIMENS AND HEALTH GUIDES (7.8%) contrast with SURGICAL AND ANA-TOMICAL TREATISES (2.9%), GENERAL TREATISES AND TEXTBOOKS (2.1%) and the *Philosophical Transactions* (0%). The latter result is perhaps not surprising given that the earliest text from the *Philosophical Transactions* dates from 1666; in GEN-ERAL TREATISES AND TEXTBOOKS, the N-forms are limited to two examples of *thyne answere* in a text from 1547 (see Table 5, which presents the results per text category and period).

Table 5. The determiners MINE/THINE and MY/THY per text category in the periods 1500–1549, 1550–1599, 1600–1649 and 1650–1700

Period	Category	MINE/THINE	MY/THY
1500–1549	TREATISES ON SPECIFIC TOPICS	1 (2%)	50 (98%)
	RECIPE COLLECTIONS AND *MATERIA MEDICA*	19 (19.4%)	79 (80.6%)
	REGIMENS AND HEALTH GUIDES	16 (9.9%)	146 (90.1%)
	SURGICAL AND ANATOMICAL TREATISES	0	2
1550–1599	TREATISES ON SPECIFIC TOPICS	16 (10.8%)	132 (89.2%)
	RECIPE COLLECTIONS AND *MATERIA MEDICA*	3 (3.3%)	89 (96.7%)
	REGIMENS AND HEALTH GUIDES	8 (7.8%)	94 (92.2%)
	SURGICAL AND ANATOMICAL TREATISES	8 (9.2%)	79 (90.8%)
1600–1649	TREATISES ON SPECIFIC TOPICS	34 (12.8%)	232 (87.2%)
	RECIPE COLLECTIONS AND *MATERIA MEDICA*	3 (4.5%)	64 (95.5%)
	REGIMENS AND HEALTH GUIDES	2 (3%)	64 (97%)
	SURGICAL AND ANATOMICAL TREATISES	1 (1.4%)	73 (98.6%)
1650–1700	TREATISES ON SPECIFIC TOPICS	6 (1.6%)	432 (98.4%)
	RECIPE COLLECTIONS AND *MATERIA MEDICA*	0	53 (100%)
	REGIMENS AND HEALTH GUIDES	0	5 (100%)
	SURGICAL AND ANATOMICAL TREATISES	2 (0.9%)	210 (99.1%)

Figure 5 illustrates the percentage (in relation to MY/THY) of MINE/THINE determiners by text category and period, excluding the *Philosophical Transactions* and GENERAL TREATISES AND TEXTBOOKS.
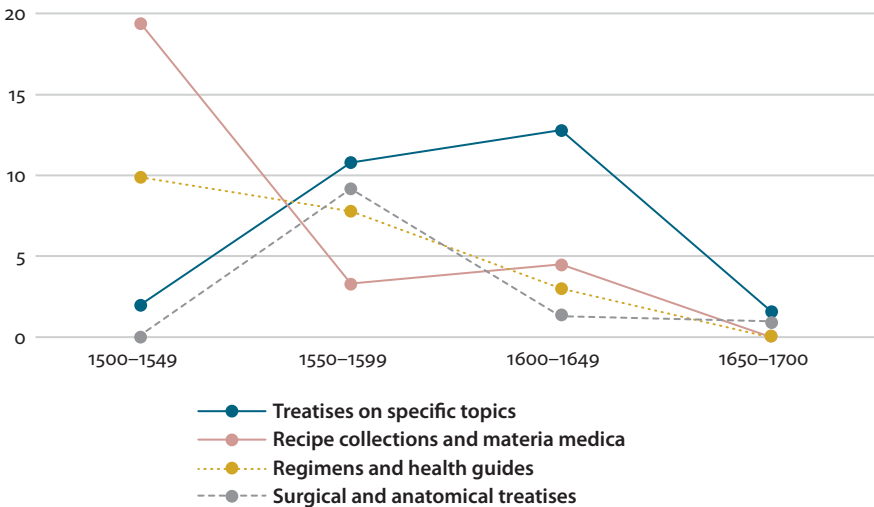
Figure 5. Percentage of MINE/THINE (in relation to MY/THY) per text category in the periods 1500–1549, 1550–1599, 1600–1649 and 1650–1700

Table 5 and Figure 5 further demonstrate the relevance of text category as regards the use of MINE/THINE. TREATISES ON SPECIFIC TOPICS stand out as the MINE/THINE forms actually increase across time (from 2% to 10.8% in the sixteenth century, rising to 12.8% between 1600 and 1649), before dropping to 1.6% in the latter half of the seventeenth century. This likely reflects the fact that this category consists of five sub-categories, which we will come back to later. RECIPE COLLECTIONS AND *MATERIA MEDICA* show a drop from 19.4% to as little as 3.3% between the first and second half of the sixteenth century, but the N-forms persist into the first half of the seventeenth century (4.5%). There is a steadier decline in the use of the N-variants in the text category REGIMENS AND HEALTH GUIDES (falling from 9.9% to 7.8% and then to 3.0%, with no N-forms after 1650). In SURGICAL AND ANATOMICAL TREATISES, MINE/THINE forms are rare in all periods other than 1550–1599 (with no occurrences in the first half of the sixteenth century); nevertheless, they survive into the final period.

If we compare the quantitative results to those of Kytö and Walker (2020), who looked at speech-related texts in the CED, we see certain similarities, although clearly there are far fewer examples of all first and second person singular determiners in EMEMT compared to the CED, a corpus of dialogic texts. In the CED, there was a decrease in N-forms before and after 1600, with a negligible number of occurrences after 1680. In the EMEMT data there is only one instance of N-forms after 1680: all other occurrences in the period between 1650 and 1700 precede the year 1680. With regard to the phonological contexts, in EMEMT, the ratio of MINE/THINE (compared to MY/THY) is 0.06% before a consonant, 5.2% before *h*, and 25.1% before a vowel sound (calculated from Figure 2). The corresponding percentages for the CED data are 0.05%, 6.8% and 25.2% respectively. As mentioned in Section 2, Kytö and Walker (2020) found constructed dialogue texts, Comedy Drama, Fiction and Handbooks, to be the text types in which the N-forms survived the longest and were the most common overall (the N-forms were barely found after 1600 in Trials and Depositions, texts based on authentic dialogues). As can be seen in Table 6, the results for the three constructed text types in the CED – allowing for the differences in periodization – are not dissimilar to those for RECIPE COLLECTIONS AND *MATERIA MEDICA* and REGIMENS AND HEALTH GUIDES, shown in Table 5. For example, the latter text category has 7.8%, 3% and 0% N-forms in the last three 50-year periods, while the text type Handbooks in the CED has 6.5%, 2.4% and 0% N-forms in the last three 40-year periods of the CED.

Table 6. Ratio of N-forms (compared to MY/THY) in *A Corpus of English Dialogues 1560–1760* (raw figures and percentages)

Text type	1560–1599	1600–1649	1649–1679
Comedy Drama	30 (3.3%)	34 (3.8%)	18 (2.7%)
Fiction	24 (5.1%)	17 (3.2%)	23 (3.1%)
Handbooks	19 (6.5%)	13 (2.4%)	0

RECIPE COLLECTIONS AND *MATERIA MEDICA* and REGIMENS AND HEALTH GUIDES were intended for a wider, popular kind of readership, perhaps carrying echoes of spoken interaction, which might then account for the similarity to the dialogic text types in the CED. What is interesting is that the survival of the N-forms in the more dialogic texts and texts intended for a wider readership seems to contrast with the finding of Schendl (1977: 188–189) that it was MY/THY that became stylistically marked as informal in the latter half of the sixteenth century.

For the microanalysis, we focus on RECIPE COLLECTIONS AND *MATERIA MEDICA* and REGIMENS AND HEALTH GUIDES and also consider the sub-categories of TREATISES ON SPECIFIC TOPICS in EMEMT. As we have shown, these exhibit the highest ratio of N-forms, and unlike for the other text categories, in which there were only 0–11 occurrences of N-forms, there may be a sufficient number of examples to allow some conclusions to be drawn as regards usage in these three text categories.

5.   The microanalysis

As demonstrated in Section 4, both RECIPE COLLECTIONS AND *MATERIA MEDICA* and REGIMENS AND HEALTH GUIDES have fairly similar patterns of decline as regards the N-variants. MINE/THINE are predominantly found in the period between 1500 and 1549, decline thereafter and are not found after 1649. TREATISES ON SPECIFIC TOPICS show a more erratic pattern, probably due to the range of sub-categories it contains. In this section, we will discuss the determiner usage in each of the three text categories in turn.

In RECIPE COLLECTIONS AND *MATERIA MEDICA*, the N-forms only occur before words starting with a vowel. Only MY/THY precede words starting with *h,* predominantly *hands* and *head*, and only two words of French origin, *herb*(*s*) and *history*. The majority of N-forms are found in the first period (1500–1549): in fact one text, the anonymous *Treasure of pore men* (1526), accounts for 17 of the 25 N-forms found in this text category overall, and are all instances of *thine eye* (or *eyes*), as in Example (1). However, *thy eye*(*s*) also occurs four times in this text. The anonymous *Newe mater* (1525) has *thine ears* twice and *thy ears* once. Other than *eyes/ears*, there are no examples of words beginning with a vowel that follow a determiner in this period.

In the period between 1550 and 1599, we find the N-forms collocating once with *infirmities* (see Example 2 from Bright's 1580 text) and twice with *own*: MY/THY are used far more before other words beginning with a vowel sound, especially in the collocations *my opinion* (also in Bright's text) and *thy use*. Fourteen examples of *thy use* are found in *Antidotarie* from 1563, written by the London surgeon Thomas Gale. These examples occur in the same formulation, always at the end of

an instruction (see Example 3): *thy use*, of course, is likely motivated by the pronunciation of the semi-vowel /j/ in *use*, and hence the non-selection of the N-form here:

(1)   [}For blered eyes.}]
      Take the ioyce of wormwode & medle it with water made of the whyte of an
      egge & putte it to **thyne eyes** & it shall put the blode & akynge away.
                                        (RECIPE COLLECTIONS AND *MATERIA MEDICA*:
                                        Anonymous, *Treasure of pore men*, 1526: F6R)[1]

(2)   For what can be more pleasant vnto thee, then the inioying of medicines for
      cure of **thine infirmities** out of **thy natiue soyle**, and countrie, **thy Fielde, thy
      Ortchard, thy Garden?**   (RECIPE COLLECTIONS AND *MATERIA MEDICA*:
                                        Timothy Bright, *Sufficiencie of English Medicines*, 1580: 7)

(3)   Put all in a leaden morter and let it be stirred two houres continually, after put
      it into a vessel of glasse. **And reserue it to thy vse**.   (RECIPE COLLECTIONS
                                        AND *MATERIA MEDICA*: Thomas Gale, *Antidotarie*, 1563: F10V)

After 1600, there are three examples of the heading "If thou wilt ouercome **thine
enimies**", but also an example of *thy enemy* from the *Secrets of Albertvs Magnvs*
published in 1617. MY/THY are used for all other words beginning with a vowel,
including *ear* and *own*. Overall, the N-forms only occur in five of the 31 texts in
the text category RECIPE COLLECTIONS AND *MATERIA MEDICA* that contain first and
second person singular determiners, and most of the examples are *thine eye(s)* in
a text from 1526.

    In REGIMENS AND HEALTH GUIDES there are only 26 examples of N-forms, again
with the majority (16 instances) occurring in the first period between 1500 and
1549, and with *thine* accounting for the majority of N-forms. Interestingly, in this
text category, the N-forms are not limited to a particular phonological context in
the period between 1500 and 1549. As illustrated in Example (4), in the earliest
text, *Gouernall of helthe* from 1506 (but first published around 1490) by the East
Anglian prior John Lydgate, N-forms occur slightly more frequently before words
with initial *h* (*humours, hour, hand(s), health*) compared to words beginning with
a vowel (*appetite, estate, old, own*), but *thine* also occurs before the word *vital* (see
Example 5).

(4)   The .viii. is that thou vse saffron in **thy mete**/ for it quyckeneth kyndly hete
      and comforteth **thy dysgestyon** and taryeth **thyne olde aege** and bryngeth in
      gladnes & letteth **thyne humors** fro rotynge and dryenge.   (REGIMENS AND
                                        HEALTH GUIDES: John Lydgate, *Gouernall of helthe*, 1506: A2V)

---

1.   The emphasis is ours in all the examples.

(5)   & yf thy metes ben to drye it throweth out **thyne vytayle vertues** & yf thy metes ben to fatte they lette **thy dysgestyon**   (REGIMENS AND HEALTH GUIDES: John Lydgate, *Gouernall of helthe*, 1506: A6V)

Similarly, in Copland's *Secrete of secretes* from 1528, we find *thine* with *enemie*s, *eyes* and *appetite*, but also with *hour*. These two texts account for all examples of *thine* in the period; one solitary example of *mine*, with *own*, is found in a text from 1547 by Andrew Boorde. While the N-forms seem to be preferred before vowels between 1500 and 1549, THY is also found with *appetite* and *enemy*. MY and THY precede *h* more often than the N-forms do, occurring with *hand* and *health* (as well as *house*(*hold*), *heart* and *head*), but not with words of French origin, such as *humours* and *hour*.

Between 1550 and 1599, MINE/THINE only precede words beginning with a vowel, but occur in this context to a greater extent than MY/THY. Gratarolo's *Health of magistrates and studentes* from 1574 has *my opinion* and *thy arms*, but *thine own* and *thine age*. Richard Mulcaster's text from 1581 has *mine opinion* but *my exercises*. The later texts from 1588 and 1595 are more consistent, and use only N-forms before words starting with a vowel (*advice, ease, eye, own*). However, after 1600, there are just two occurrences of N-forms: *thine own* is found in Bacon's 1625 text and *in mine opinion* is found in Venner's text from 1620; however, later in the text Venner uses *in my opinion* three times (twice as a parenthetical comment). Only the MY/THY forms occur before a vowel in four other texts from this period, and in two texts before *h*. Overall in REGIMENS AND HEALTH GUIDES, the N-forms are found in nine of the 22 texts exhibiting first and second person determiners.

TREATISES ON SPECIFIC TOPICS attest N-forms in each of the four periods, occurring in 23 of the 59 texts with first and second person determiners, but are distributed unevenly across the five sub-categories of PLAGUE, SPECIFIC DISEASES, SPECIFIC THERAPEUTIC SUBSTANCES, SPECIFIC METHODS, and MIDWIFERY AND CHILDREN'S DISEASES. Plague texts contain 18 examples of MINE/THINE in the five texts that contain first and second person singular determiners. In Moulton's 1539 text we find *mine own* and *thy arm*, but only MY/THY are found before words beginning with vowels and *h* (none of the latter are words of French origin). In Brasbridge's *Poore mans iewel* from 1578, we find *thine hands* but *thy heart* and *my opinion*. A text from 1603, *Plagves infection*, by London clergyman James Balmford, accounts for 12 of the 18 N-forms in this sub-category. This text is presented as a dialogue, similar to the Handbooks sampled in the *Corpus of English Dialogues*, as illustrated in Example (6). Only N-forms are used before a vowel and before *h*, with the exception of the Bible quote "the Lord is **my hope**".

(6)   [}Profess.}] […] If the Plague be contagious, why is not one infected as well as another? […] I (and a great number besides me, who haue done as much) had neuer the plague yet, and trust neuer shal, so long as I haue a strong faith in God: for is it not written, Thou shalt not be afraid of the pestilence, for thousands shal fal besides thee, yet it shall not come neare thee; for thou hast said, The Lord is **my hope**.
      [}Preacher.}] […] Is <u>thine eye</u> euill because God is good? Wilt thou by **thy bloody errour** poison other, because God hath glorified his speciall prouidence ouer thee? Is this **thy thankfulnesse** for so great deliuerance, to obscure Gods prouidence by attributing <u>thine escape</u> to this, that the plague is not infectiue?
      (TREATISES ON SPECIFIC TOPICS:
      James Balmford, *Plagves Infection*, 1603: 44–45)

Another text from 1603 has *mine owne opinion*, but otherwise uses MY before vowels; as late as 1665, in the text *Food and Physick*, we find *thy angels* but *thine anger, arm* and *ear*. Texts in the sub-category SPECIFIC DISEASES contain 16 N-forms: 10 of these occur only before vowels, especially with the word *own*, in three texts from the second half of the sixteenth century (dating from 1552, 1596 and 1599), and each of these texts only has one occurrence of MY before a vowel. After 1600, there are six examples of MINE before vowels in texts from 1602, 1616 and 1621, as in "**mine eyes** are sorrowfull through **mine affliction**" (Robert Burton, *Anatomy of Melancholy*, 1621: 4), but with the exception of this text, the MY/THY forms are used to the same extent before vowels. The only N-forms found in the sub-category SPECIFIC METHODS occur in *The Pisse-prophet* from 1637, illustrated in Example (7). This is "a polemical book against the use of urine for diagnoses", written by Thomas Brian, a physician practicing in London and Colchester (EMEMT Catalogue): MINE collocates with *advice, own* and *opinion*; other words beginning with a vowel are preceded by MY (which also collocates twice with *opinion*).

(7)   This Nurse tells me indeed (but if she had not told me, I should have perceived it) that she came to heare <u>mine opinion</u> of the water, and that the sicke party would send to me againe, when she had heard (she should have said, If she liked it) what was <u>mine opinion</u>, and I thinke I have fitted her, and now tell the Nurse, that it had been fitter they had sent for some present remedies, than to know the disease, or to heare <u>mine opinion</u> of it, so I hast her away with her oracle, bid her be sure to deliver it so as I had told her, and tell (but not to the sicke party) to some of her neere friends, how dangerously she is sicke, and that I would (if she thinke good to make use of me) use the best meanes to recover her that Art could lead me unto; […]   (TREATISES ON SPECIFIC TOPICS:
      Thomas Brian, *The Pisse-prophet*, 1637: 23–24)

Specific therapeutic substances attest 11 n-forms in total. In the 1580s only mine occurs before a vowel, collocating with *opinion* and *own*. Between 1600 and 1649, there is an equal number of examples of mine and my before vowels, although the word *own* only occurs after mine. In the sub-category midwifery and children's diseases, there are only four examples of n-forms, which co-occur with *eyes* and *own*. In 1612, there is only one example of a determiner before a vowel, *mine own*, whereas in later texts, both n-forms and my/thy occur before vowels. Nicholas Culpeper's 1651 text *Directory for Midwives* has both mine and my before *own*, but my/thy before other vowels, and in the anonymous 1684 text, *Aristoteles master-piece,* both *mine eyes* and *thy eyes* occur.

Overall in EMEMT, the words that collocate most frequently with the n-forms are *eye/eyes* (24 examples in eight texts from 1526 to 1684), *opinion* (21 examples in 10 texts from 1562 to 1637), and *own* (32 examples in 23 texts from 1506 to 1651).

6.   Summary and concluding remarks

The present study turned to an important resource in the study of the development of the possessive determiner forms my, mine, thy, and thine, namely the corpus of *Early Modern English Medical Texts* (EMEMT), covering the period from 1500 to 1700. The rich representation of text categories within this specialized area of writing, extending from learned professional treatises to popular recipe collections, health guides and other documents intended for layman readership, made this resource of particular interest to the study of change in the occurrence of the forms in the Early Modern English period.

The linguistic factors considered in this study included the phonological environment: mine/thine forms preceding vowels accounted for most of the instances attested. The n-forms had a healthy majority over my/thy in this context before 1550, but dropped to around 8% in the following century, with the final death throes occurring after 1650. Mine/thine preceding *h* became rare after 1550, with no survivors at all after 1650.

The text categories distinguished in the corpus appear to be an important factor influencing the distribution of the forms. Three text categories promoted the use of the n-forms, treatises on specific topics, recipe collections and *materia medica*, and regimens and health guides, whereas they were rare in surgical and anatomical treatises and in general treatises and textbooks, and non-existent in the *Philosophical Transactions*.

Comparisons with the data drawn from the *Corpus of English Dialogues 1560–1760* (CED) showed that 1680 marked a watershed in both corpora: the number of occurrences with the n-forms in the CED was negligible after 1680, and only

one instance was recorded in EMEMT after this year. Both corpora showed con-
curring developments in the phonological contexts studied. Comparisons with
the CED pointed to a factor likely to help account for the patterns of use attested
in EMEMT. In the CED, the constructed text types of Comedy Drama, Fiction
and Handbooks helped the N-forms survive longer, after they had drawn their
last gasp in Depositions and Trials around 1600: the popular text categories rep-
resented by RECIPE COLLECTIONS AND MATERIA MEDICA, REGIMENS AND HEALTH
GUIDES, and TREATISES ON SPECIFIC TOPICS in EMEMT could be assumed to
share some affinities with the dialogues in the constructed text types in the CED.
Nevertheless, this seems to conflict with Schendl's (1997) finding for the *Helsinki
Corpus* data, in which the more informal contexts appear to have promoted the
MY/THY forms. Taking into account the age variable might offer further insights
on this apparent contradiction, as Schendl (1997: 186) illustrated how this might
play a role in his data, but did not examine the factor systematically. As mentioned
in Section 2, we have not looked at sociolinguistic variables in this study since in-
formation on age is not always available for our data. In other aspects our findings
tend to be more in line with those of earlier researchers, such as Nevalainen and
Raumolin-Brunberg (2003).

The microanalysis focused on the three text categories sustaining the use of the
N-forms in EMEMT, and highlighted the characteristic collocations in which the
MY/THY and MINE/THINE forms tended to occur. It is important to comment on the
dispersion of MINE/THINE across the texts and text categories. In RECIPE COLLEC-
TIONS AND MATERIA MEDICA one early text accounted for most of the instances of
N-forms, with only 16% of the texts (5 of 31) containing N-forms at all; thus the re-
sults for this text category mostly reflect usage in the popular remedy book *Treasure
of pore men*. In REGIMENS AND HEALTH GUIDES, dispersion of the N-forms across
texts (9 of 22) is much greater (41%), but, again an early text, Lydgate's *Gouernall
of helthe* (1506), was responsible for many of the N-forms (10 of 26). The N-forms
have a similar dispersion across texts in TREATISES ON SPECIFIC TOPICS (23 of 59, or
39%), but this varied widely across the sub-categories. In the PLAGUE sub-category,
all but one of the texts contained N-forms, yet Balmford's *Plagves infection* from
1603 – a dialogue similar to those in Handbooks in the CED – accounted for
two-thirds of MINE/THINE examples. Furthermore, of the seventeen texts in the
sub-category SPECIFIC METHODS, the 1637 text *The Pisse-prophet* accounted for all
the N-forms: the speech-like polemic argumentation in this text may have encour-
aged the use of these forms.

Finally, the study pointed to the importance of specialized historical corpora,
such as EMEMT, which allow researchers to probe deeper into the change in pat-
terns of usage in the history of English. The picture we obtain of standardization
phenomena such as the decline and final decease of the MINE and THINE forms is

bound to remain defective or at least incomplete if data from specialized corpora targeting various areas of speech-related and non-speech-related writing is not taken into account.

References

Electronic resources

Corpus of Early English Correspondence (CEEC). 1998. Compiled by Terttu Nevalainen, Helena Raumolin-Brunberg, Jukka Keränen, Minna Nevala, Arja Nurmi and Minna Palander-Collin. Department of Modern Languages, University of Helsinki.

A Corpus of English Dialogues 1560–1760 (CED). 2006. Compiled by Merja Kytö (Uppsala University) and Jonathan Culpeper (Lancaster University).

Early Modern English Medical Texts (EMEMT). 2010. Compiled by Irma Taavitsainen, Päivi Pahta, Turo Hiltunen, Martti Mäkinen, Ville Marttila, Maura Ratia, Carla Suhr and Jukka Tyrkkö. Software by Raymond Hickey. See Taavitsainen and Pahta (2010) below.

Helsinki Corpus = The Helsinki Corpus of English Texts. 1991. Compiled by Matti Rissanen (Project leader), Merja Kytö (Project secretary); Leena Kahlas-Tarkka, Matti Kilpiö (Old English); Saara Nevanlinna, Irma Taavitsainen (Middle English); Terttu Nevalainen, Helena Raumolin-Brunberg (Early Modern English). Department of Modern Languages, University of Helsinki.

Secondary sources

Barber, Charles. 1976. *Early Modern English*. 1st edition. London: André Deutsch.

Barber, Charles. 1997. *Early Modern English*. 2nd edition. Edinburgh: Edinburgh University Press.

Busse, Ulrich. 2002. *Linguistic Variation in the Shakespeare Corpus: Morpho-syntactic Variability of Second Person Pronouns*. Pragmatics and Beyond New Series 106. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.106

Hilpert, Martin. 2013. *Constructional Change in English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139004206

Jespersen, Otto. 1927. *A Modern English Grammar on Historical Principles*. Volume II. Heidelberg: Winter.

Jucker, Andreas H., and Irma Taavitsainen. 2014. "Diachronic Corpus Pragmatics: Intersections and Interactions." In *Diachronic Corpus Pragmatics*, ed. by Irma Taavitsainen, Andreas H. Jucker, and Jukka Tuominen, 3–26. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.243.03juc

Kytö, Merja, and Terry Walker. 2020. "A Standardization Process in Its Final Stages: *Mine* and *Thine* in *A Corpus of English Dialogues 1560–1760*." In *Standardisation and Change in Early Modern English: Empirical Approaches*, ed. by Laura Esteban-Segura and Javier Calle-Martin. Special Issue of *International Journal of English Studies* 20(2): 95–116. https://doi.org/10.6018/ijes.369761

Marttila, Ville. 2010. "Recipe Collections and *Materia Medica*." In *Early Modern English Medical Texts. Corpus Description and Studies*, ed. by Irma Taavitsainen and Päivi Pahta, 101–110. Amsterdam: John Benjamins. https://doi.org/10.1075/z.160.04cat3

Nevalainen, Terttu, and Helena Raumolin-Brunberg. 2003. *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. London: Longman.

Pahta, Päivi, and Irma Taavitsainen. 2010. "Introducing Early Modern English Medical Texts." In *Early Modern English Medical Texts. Corpus Description and Studies*, ed. by Irma Taavitsainen and Päivi Pahta, 1–8. Amsterdam: John Benjamins. https://doi.org/10.1075/z.160.00toc

Raumolin-Brunberg, Helena, and Terttu Nevalainen. 2007. "From *Mine* to *My* and *Thine* to *Thy*: Loss of the Nasal in the First and Second Person Possessives." In *Tracing English through Time: Explorations in Language Variation*, ed. by Ute Smit, Stefan Dollinger, Julia Hüttner, Gunther Kaltenböck, and Ursula Lutzky, 303–314. Vienna: Braumüller.

Rühlemann, Christoph, and Karin Aijmer. 2015. "Corpus Pragmatics: Laying the Foundations." In *Corpus Pragmatics: A Handbook*, ed. by Karin Aijmer and Christoph Rühlemann, 1–26. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139057493.001

Schendl, Herbert. 1997. "Morphological Variation and Change in Early Modern English: *My/Mine, Thy/Thine*." In *Language History and Linguistic Modelling*, ed. by Raymond Hickey and Stanisław Puppel, 179–191. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110820751.179

Taavitsainen, Irma, and Päivi Pahta (eds.). 2010. *Early Modern English Medical Texts. Corpus Description and Studies*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.160

Taavitsainen, Irma, and Jukka Tyrkkö. 2010. "The Field of Medical Writing with Fuzzy Edges." In *Early Modern English Medical Texts. Corpus Description and Studies*, ed. by Irma Taavitsainen and Päivi Pahta, 57–62. Amsterdam: John Benjamins. https://doi.org/10.1075/z.160.04taa

Walker, Terry. 2007. Thou *and* You *in Early Modern English Dialogues: Trials, Depositions, and Drama Comedy*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.158

Towards a local grammar of stance expression in Late Modern English medical writing

Turo Hiltunen

University of Helsinki

This chapter investigates the expression of stance in eighteenth-century medical English, based on data from the LMEMT corpus. The focus is on three types of stance *that*-clauses, and the aim of the analysis is to document the frequency trends and patterns of variation in the data, compare them with those identified for medical texts from previous centuries, and provide a functional description of their use within the framework of local grammar. With respect to the frequencies of stance *that*-clauses across different sub-registers, the eighteenth century represents a period of stylistic continuation. The degree of phraseological variation is considerable, and the analysis demonstrates the usefulness of a local grammar approach to account for this variation.

Keywords: stance marking, sub-register, local grammar, eighteenth-century medical writing

1.   Introduction

The pragmatic function of *stance marking* – the linguistic expression of a writer's personal attitudes, evaluations and commitments – is a popular topic in studies on Present-day academic English, and work has also been carried out to describe historical patterns across different registers. Previous work has likewise established stance as a useful concept for a variety of purposes: on the one hand, it provides a meaningful way of describing differences between registers of writing, and as such may be useful in teaching tacit rules of academic argumentation to students (e.g. Hyland and Sancho Guinda 2012). On the other hand, historical studies have documented general changes, such as a general increase in stance expressions especially in the twentieth century, and variation in the rates of use of specific stance markers (Biber 2004).

Stance marking is also clearly relevant to the analysis of medical discourse and medical writing, both synchronically and diachronically. Stance expressions

are typically linked with functions like indicating the source of information and assessing its certainty, which are obviously crucial when dealing with issues of health and sickness.

This chapter sets out to investigate one aspect of stance expression in medical writing 1700–1800. This century represents an important transition period in the history of English medicine, characterised by a growth in the volume of publication and the range of topics, the consolidation of journal publications, and the spread of philanthropic ideas and an increased interest in public health (see further Taavitsainen, Jones, and Hiltunen 2019). Despite these developments, eighteenth-century medical argumentation has attracted comparatively little scholarly attention, and stance marking in medical texts has likewise not been systematically investigated. Using the recently published *Late Modern English Medical Texts* corpus (LMEMT; Taavitsainen et al. 2019) as the source of data, the present study takes the first step towards this direction.

This chapter adopts a corpus-pragmatic approach, integrating "vertical" (concordance-based) and "horizontal" (close-reading) methodologies (Rühlemann and Aijmer 2015: 3).[1] These methodologies are closely linked to the two aims of this study: documenting patterns of variation and change of stance marking in medical writing, and providing a functional description that is nuanced, methodologically sound, and has optimal explanatory power. As for the former aim, the data on stance-expressing declarative *that*-clauses are analysed quantitatively, and the findings are compared to previous analyses on Early Modern English medical data (Gray, Biber, and Hiltunen 2011). And as for the latter, the concept of local grammar (Hunston and Sinclair 2000) will be used for a close analysis of the data, in order to determine the phraseological and discourse-pragmatic factors accounting for the patterns observed in the corpus data.

2.   Theoretical background

2.1   Stance

In its most general sense, the term *stance* encompasses linguistic resources that speakers/writers may use to mark their personal attitude towards or an assessment of their statements, with respect to such parameters as certainty and reliability of

---

1.   This terminology runs parallel to Tognini-Bonelli's (2001) distinction between two ways of reading: an individual *text* is read "horizontally" for content, whereas a *corpus* is read "vertically" for recurrent patterning (2001: 2–3). For methodological overviews of historical/diachronic corpus pragmatics, see also Kohnen (2009) and Jucker and Taavitsainen 2014).

information, the positivity/negativity of their value, or their importance or relevance (e.g. Gray and Biber 2012). However, like many widely applied concepts in linguistics, this term is also used in somewhat varying ways in different areas (for a conceptual overview, see e.g. Jaffe 2009: 4–7), and its meaning also partially overlaps with that of other, related terms such as *evaluation, evidentiality, affect, point-of-view, positioning*, and *appraisal* (Hyland and Sancho Guinda 2012). Stance can likewise be, and has been, operationalised for investigation in multiple ways. Specifically for diachronic studies, the three main approaches according to Landert (2019: 171) are

I.   to investigate individual markers in isolation
II. to focus on a number of conventionalised lexicogrammatical patterns which are retrieved semi-automatically, and
III. to carry out comprehensive qualitative analysis of all stance devices through close reading.[2]

Landert (2019) also observes that each approach has its own advantages and limitations from the perspective of the methodology of corpus linguistics and historical pragmatics: while (I) is usually the most straightforward one to implement, it is limited in generalisability due to its narrow focus. By contrast, (III) offers optimal precision and recall, but the analysis is necessarily limited to a small number of texts, due to the fact that the analysis is very labour-intensive. From this perspective, approach (II), associated with the work of Biber (2004), would appear to strike a balance between feasibility and generalisability in that it offers a reasonable recall by covering a number of potential markers, while still making use of the speed and convenience of automated text analysis. For this reason, it is also the approach that is adopted for the quantitative part of this study.

The premise of approach (II) is the observation that speakers' and writers' expression of evaluations, values, judgements and feelings is routinely done through using specific lexicogrammatical resources that are available for this purpose. In other words, certain vocabulary items and grammatical constructions are treated as stance markers due to their regular association with this discourse function, and after such markers have been identified, it is possible to analyse them quantitatively, and for example compare different registers with respect to how often stance is expressed in them and what specific stance markers are used (e.g. Biber et al. 1999: 969–970; 980). This approach has been particularly influential in the study of present-day academic discourse, where it is frequently seen as one aspect

---

2.   In addition, a fourth, promising approach is proposed by Landert (2019: 171). This approach focuses on identifying "text passages with a high-density of stance markers" with the help of a number of "reliable indicators of stance".

of *metadiscourse* (e.g. Hyland 2005a, 2005b; McGrath and Kuteeva 2012). In addition to this work characterising stancetaking on a general level, there are many other studies investigating the use of specific linguistic features associated with stance, such as personal pronouns, *that*-clause constructions, adjectives, adverbs and adverbials (e.g. Kuo 1999; Charles 2007; Hyland and Tse 2005b; Harwood 2006; Finegan 2009; Hiltunen 2010; Goźdź-Roszkowski 2017, 2019). These studies are also relevant to the present investigation, yet the period and differences in genre, discipline, and the sociohistorical context need to be taken into account.

Stance marking has been investigated from a historical perspective both across registers of English more broadly (Biber and Finegan 1989; Biber 2004), and specifically applied to registers and sub-registers of medical writing in the early modern period and in the second part of the twentieth century (Gray, Biber, and Hiltunen 2011; Hyland and Jiang 2016). Particularly relevant to the present investigation is Gray, Biber, and Hiltunen's (2011) study of stance expressions across different categories of medical texts in the sixteenth and seventeenth centuries, with the following observations:

–   The overall frequency of stance marking varies markedly between medical genres, as do the relative frequencies of different lexicogrammatical markers.[3]
–   The dominant stance meaning attested in early modern English medical writing was found to be the expression of epistemic certainty.
–   Recipe collections and *materia medica* tended to contain relatively few stance expressions.
–   Articles from the journal *Philosophical Transactions*, particularly book reviews, exhibited a high frequency of *that*-complement clauses licensed by verbs, which reflects the communicative purposes of this emerging medium of science communication: mediating both original research findings and ideas expressed in contemporary medical books of the community of physicians and natural philosophers of the early Royal Society.

These findings provide a convenient baseline that the quantitative findings of the present study can be meaningfully compared to.

In corpus-pragmatic terms, approach (II) represents a "vertical reading" methodology (Rühlemann and Aijmer 2015), where insights are based on corpus frequencies and co-occurrence patterns, rather than the consideration of individual examples. Despite the many advantages of this approach, its most obvious limitation is precisely the lack of attention to individual pragmatic meanings of specific

---

3.   Note that Gray, Biber, and Hiltunen (2011) investigated not only *that*-clauses but also modals and adverbs.

texts and utterances. In other words, due to its focus on general trends across a large volume of texts and the semi-automatic treatment of data, some potentially meaningful distinctions are necessarily glossed over, both at the level of how stance features are classified, and what functions and local meanings they express in individual texts. To illustrate this issue, we can look at the verb *doubt*, which in Gray, Biber, and Hiltunen (2011: 256) is classified as a verb indicating *likelihood*. Yet we can observe in the data that this verb is predominantly used in constructions which in fact signal a high degree of epistemic *certainty*, such as Example (1):[4]

(1)   nor is it to be **doubted**, *that when the Nerves of the Stomach are, for certain causes, affected in a manner somewhat different, an indifference for food, a weak Digestion, a languor and coldness, a slow Pulse, and wasting, may be the consequences* […]

(GEN, Anonymous, *A dissertation upon the nervous* system, 1780)[5]

The point here is not to criticise or contest the classification of individual items like *doubt*, as these are unlikely to have a major influence on the main quantitative patterns, given the large number of items included in the analysis. However, the broader issue is that stance meanings are not just conveyed by individual lexical items and their colligations, but other phraseological features (e.g. negators), and configurations of role relationships of social actors in the text clearly contribute to their meaning. While previous frequency-based analyses of stance are also complemented with qualitative observations, the present study addresses this issue systematically using the notion of *local grammar*, described in the following section.

2.2   Local grammar

As a concept, local grammar emerged around the turn of the previous century in the studies of scholars of the Birmingham school of corpus linguistics, following the work of John Sinclair (e.g. Sinclair 1991).[6] This body of research adopted a corpus-driven approach to describing the grammar of English, and its outcomes are summarised in a number of publications (e.g. Hunston and Sinclair 2000; Barnbrook 2002; Hunston 2003). The concept of local grammar has more recently

---

4.   We can also note that the noun *doubt* is classified as a *certainty* noun (Gray, Biber, and Hiltunen 2011: 256), due to its strong association with expressions like *there can be no/little doubt that*. For further discussion on this issue, see Landert (2019: 179–80).

5.   The short titles after the quoted examples refer to the LMEMT corpus (Taavitsainen et al. 2019), where the full bibliographical details of the texts are available. Emphasis in all examples is mine.

6.   See further McEnery and Hardie (2012: 79–81; 122).

gained momentum in several works where it has been used to shed light on the phraseological realisations of a variety of discursive acts (e.g. Hunston and Su 2017; Su and Wei 2018). Historical and conceptual overviews of the developments in the analysis of local grammars are provided in Bednarek (2008: 66–69) and Su (2018: 85–87), the latter focusing particularly on speech acts.

In this study, I follow the definition of Hunston and Su (2017: 571–2), which treats *local grammar* as a grammar of discourse function, either defined in relation to specific performative speech acts such as apology and thanking (Su and Wei 2018; Su 2018), or applied to more general and diverse discourse acts like expressing affect (Bednarek 2008). The present study represents the latter perspective, in that the discourse function in focus is what in previous research has variously been termed *stance* or *evaluation*: "expression of an attitude towards an entity" (Hunston and Su 2017: 572). More specifically, the focus is on a set of related constructions involving declarative content clauses: those licensed by verbs, nouns and adjectives, illustrated in Examples (2–4). As previously mentioned, these constructions have been identified both as being important in academic texts due to their use in expressing and reporting "knowledge claims", as well as their potential for indicating stance and evaluation (e.g. Hyland and Tse 2005a, 2005b; Charles 2006, 2007; Hiltunen 2010). Due to this association, they will here be referred to as "stance *that*-clauses".

(2) These Facts; and Observations, to mention no more, are I think sufficient to demonstrate, **that** there is no such Intercourse and Communication between the Glands and Membranous Cells and the Blood, as is commonly Imagined; […]   (GEN, Hooke, *An Essay On Physick*, 1734)

(3) This to me is an evident Proof **that** this Gentleman was entirely lost by this Neglect, the Disease being too violent, and too sudden to be conquer'd by internal Medicines; […]
       (SP-METH, Catherwood, *A New Method Of Curing The Apoplexy*, 1715)

(4) I think it's evident, **that** Dr. Chamberlen's main design, in his preface, and notes upon Mr. Mauriceau's book, was to puzzle, rather than instruct the Midwomen or Mid- men, to persuade them, that he, and the rest of the Nostrum-mongers, could serve them better than any other artists […]
                      (SP-MW, Chapman, *A Reply To Mr Douglass*, 1737)

Local grammar may thus provide a more nuanced description of specific discourse functions, and as such may serve as a useful complement to the quantitative register-based approach. From a corpus-pragmatic perspective, it represents the process of "horizontal reading" (Rühlemann and Aijmer 2015: 3), which can be used to describe in detail the contextual features that are relevant in a specific communicative situation. The qualitative nature of this approach also means that it

is only manageable for moderately-sized datasets, and therefore cannot be applied to the complete set of nearly 8,000 stance *that*-clauses in the corpus. In this study, the local grammar analysis is applied to a judiciously selected random sample of instances (see Section 5.1).

3.   Material and methods

3.1   Corpus

The data for this study comes from the *Late Modern English Medical Texts* corpus (LMEMT) (Taavitsainen et al. 2019). The corpus is divided into seven text categories, which enable the study of sub-register differences within medical writing.[7] The categories, which are based on text-external factors (Hiltunen and Taavitsainen 2019), are:

1. *General treatises and textbooks* (GEN)
2. *Specific treatises* (SPEC) (includes four sub-categories):
    – *Diseases* (SP-DIS)
    – *Methods* (SP-METH)
    – *Therapeutic substances* (SP-TH)
    – *Midwifery* (SP-MID)
3. *Medical recipe collections* (RECIP)
4. *Regimens and health guides* (REG)
5. *Surgical and anatomical texts* (SURG)
6. *Public health* (PUB)
7. *Scientific periodicals* (PER) (includes two sub-categories):
    – *Philosophical Transactions* (SC-PER-PT)
    – *Edinburgh Medical Journal* (SC-PER-EMJ)
8. *The Gentleman's Magazine* (GM)

The category *General treatises and textbooks* contains extracts from medical treatises and textbooks covering (or purporting to cover) all fields of medicine.[8] With four sub-categories, *Specific treatises* is a broadly similar category as far as textual

---

7.   The categories of LMEMT are treated here as sub-registers in the sense of Biber and Conrad (2019), but this is not the only way to use them. As discussed in Hiltunen and Taavitsainen (2019), the categories are identified based on the consideration of various disciplinary factors and medical history, and do not directly correspond to any text-linguistic or discourse-analytical labels.

8.   Such books may extend to hundreds of pages, and are represented in the corpus by shorter extracts.

characteristics go, but the scope of these texts is different as their focus is on specific illnesses and methods of treatment. However, more radical differences in communicative purpose can be observed in the remaining categories. *Medical recipe collections* focus on the preparation and administration of remedies. *Regimens and health guides* discuss the preservation of health, and *Surgical and anatomical texts* provide instructions on how to perform surgical operations. *Public health* texts contain texts anticipating the public health movement of the nineteenth century. The remaining two categories represent the genre of the journal, which had emerged in the previous century with the establishment of the *Philosophical Transactions of the Royal Society* (PT); along with the PT, the LMEMT includes another scientific periodical, the *Edinburgh Medical Journal* (EMJ) as well as a general periodical, *The Gentleman's Magazine* (GM). More detailed descriptions of the text categories, as well as their development specifically in the eighteenth century, are provided in the relevant sections of Taavitsainen and Hiltunen (2019).

The distribution of texts and their word counts is shown as a scatterplot in Figure 1, where each dot corresponds to one text. As can be seen, the word counts of individual texts vary somewhat, due to the fact that short texts were included in their entirety whereas longer texts are represented by roughly 10,000-word samples.
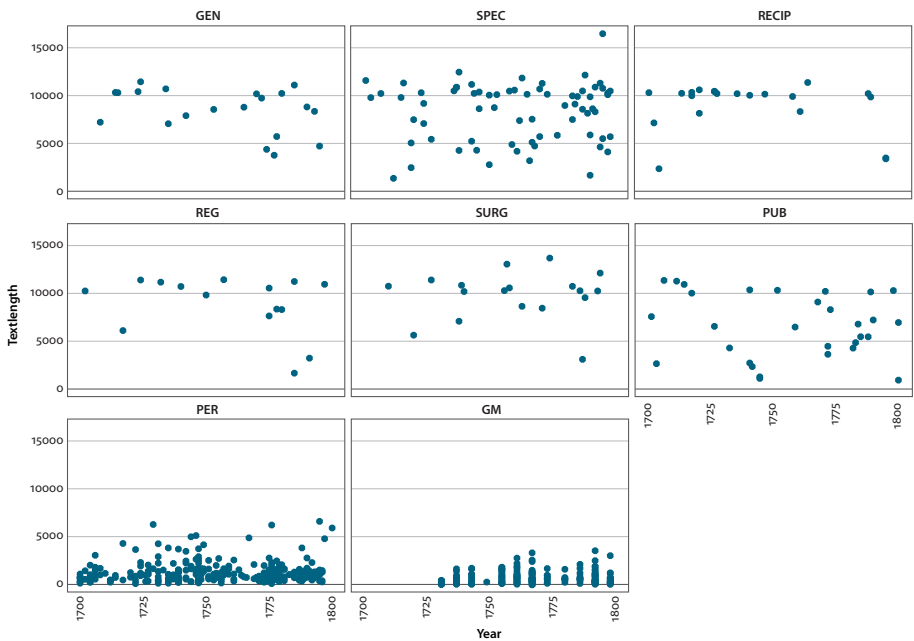
Figure 1. Scatterplot showing text categories and word counts of individual texts in the LMEMT corpus

As far as text length is concerned, the figure highlights the fact that journal articles (PER and GM) are different from the majority of texts in the other categories in being considerably shorter. While this difference does not preclude cross-category comparisons, it might be a cause of some of the differences observed in the data. The other categories included in the study do not display major differences in the mean text length.

The corpus was POS-tagged (part-of-speech) for this study using the CLAWS tagger (Garside and Smith 1997).[9]

3.2 Method of analysis

3.2.1 *Retrieval and frequency analysis*
All instances of *that*-clauses were exhaustively retrieved from the corpus, false hits were removed,[10] and the licensing word and its word class was identified with the help of POS-tags. To compare the quantitative results to previous research, particularly Gray, Biber, and Hiltunen (2011) on medical writing from two previous centuries (1500–1700), the frequencies of the three constructions were also determined across the categories listed above. The results of this part of the analysis are reported below in Section 4.

3.2.2 *Local grammar analysis*
The method of analysing local grammar of stance is adapted from Hunston and Su (2017), who suggest the following workflow:

I. Identifying a recurring sequence
II. Mapping functional labels onto the sequence, by specifying
   a. the function
   b. the realisation in lexicogrammatical terms
   c. functional labels for annotating representative examples
      (Hunston and Su 2017: 571)

---

9. A spelling-normalised version of the LMEMT corpus has not been released, and the spelling of the texts was not normalised prior to tagging. While variant spellings are a potential source of tagging errors, this is less of a problem for eighteenth-century texts than for older texts (Hiltunen and Tyrkkö 2013), and the tagged corpus was deemed sufficiently accurate for the present purpose through manual inspection.

10. This involved weeding out tokens of *that* functioning as a relativiser or a demonstrative, as well as actual content clauses in other sentential functions, such as subjects and predicative complements.

The recurring sequence in focus is the declarative content clause licensed either by verbs, nouns or adjectives (Examples 2–4 above); the linkage of these constructions with stance expression is well-established in previous work, as discussed above (see also Menzel, this volume). These three construction types were considered in the frequency analysis, while the analysis of local grammar was limited to verb-licensed stance *that*-clauses, whose form is the most frequent type of stance *that*-clause in the data (5,783 tokens).

To explore the phraseological patterning of verb-licensed clauses, a frequency-ranked list of licensing verbs was first created. With the help of this list, a representative sample was extracted from the entire dataset, which was annotated with respect to three functional elements adapted from Hunston and Su (2017: 582), illustrated in Table 1. With the help of these elements, a number of recurrent phraseological patters were identified and described, which form the empirically-determined local grammar, summarised in Section 5. The analysis is complemented by a brief illustration of how select elements function rhetorically as part of the argument in specific medical texts.

Table 1. Meaning elements considered in local grammar analysis

Element	Example	Explanation
TARGET	*Some suppose,* **that the effect of the admission of this air into the blood is a fermentation**	Claim, idea or statement being evaluated
SOURCE	*I have asserted, that my Lotion is a Specific for these Eruptions*	Person responsible for the evaluation (here: writer of text)
PROCESS	*Thus we may* **believe** *that this cold Humour is also sent thither from the salivary Sources*	Type of process involved in the evaluation (here: cognitive)

Finally, Section 6 provides a synthesis of these two stages of analysis, considering how the sets of local grammar can be used to explain and investigate the trends observed in frequency data.

4.   Frequency analysis

The quantitative analysis of data indicates, firstly, that stance is expressed at varying rates in different categories of medical writing. This can be observed in Figure 2, which shows the mean frequencies of all stance *that*-clauses per category; the categories have been ranked from the lowest to the highest ratio of stance marking. As can be seen, the mean text frequency is lowest in *Medical recipe collections*, which matches the finding of Gray, Biber, and Hiltunen (2011) on Early Modern English

data, suggesting that this stylistic characteristic of medical recipes has remained constant from the sixteenth to the end of the eighteenth century. Another category characterised by a comparatively low ratio of stance marking through *that*-clauses is *Regimens and health guides*. The other categories appear relatively close to one another, but the wide confidence intervals plotted in Figure 2 also suggest that there is considerable variation within individual categories.
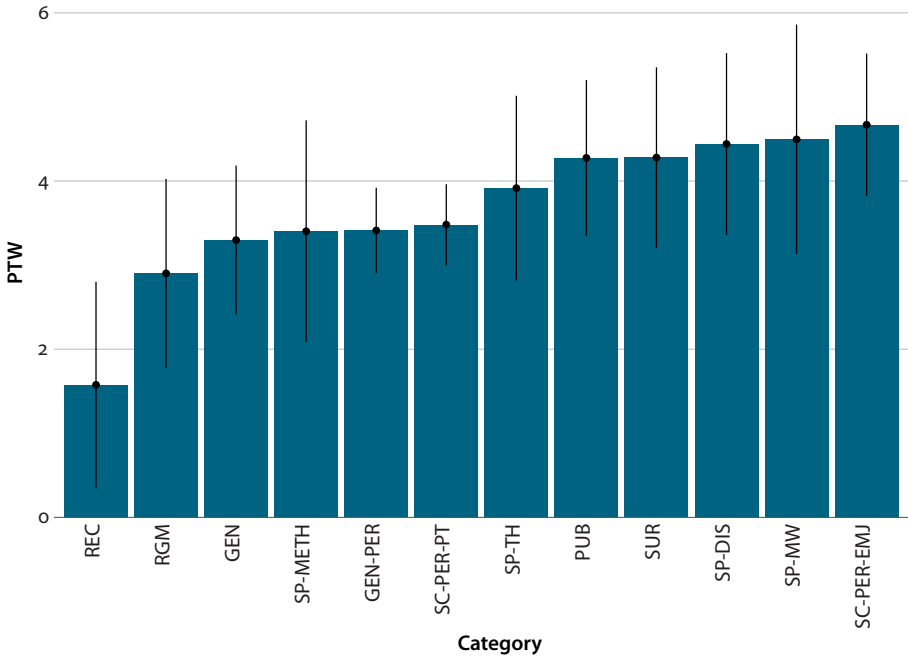
Figure 2. Mean text frequency of all stance *that*-clauses per category (the error bar indicates a 95% confidence interval)

Interestingly, we can no longer observe major differences between *Periodicals* and the other categories, which partly runs counter to Gray, Biber, and Hiltunen's (2011) findings for the previous centuries. The category with the highest rate of stance *that*-clauses is still *Scientific Periodicals: Edinburgh Medical Journal*, but it is closely followed by four other categories (*Specific treatises: Midwifery, Specific treatises: Diseases, Surgical and anatomical texts* and *Public health*); in addition, the other two categories of *Periodicals* (*Philosophical Transactions* and *The Gentleman's Magazine*) clearly exhibit lower frequencies. The reasons for this apparent convergence are not entirely clear at this point. It is possible that the other categories have become more similar to periodicals in this respect, or that the central tendencies observed for periodicals are influenced by discursive and stylistic changes in this sub-register

(see in particular Bazerman 1988; Atkinson 1999) or differences in sampling,[11] but further investigation is clearly needed.

Figure 3 shows that the distributions of different types of stance *that*-clauses are largely uniform across the categories. Verb-licensed stance *that*-clauses are by far the most frequent type of clause, yet there is some variation in their relative frequencies. Depending on the category, they make up between 74% and 88% of all stance complement clauses. Adjective- and noun-licensed clauses are much less frequent across the board. Figure 3 also provides more nuance to the overall frequency patterns discussed above. For example, the overall high frequency of stance complement clauses in *Scientific Periodicals: EMJ* observed in Figure 2 is predominantly due to verb-licensed clauses, whereas noun-licensed clauses are particularly infrequent in *Medical recipe collections* (9 occurrences, accounting for 3% of the total). The distribution of types of stance *that*-clauses, too, is broadly similar to the findings reported in Gray, Biber, and Hiltunen (2011: 235).
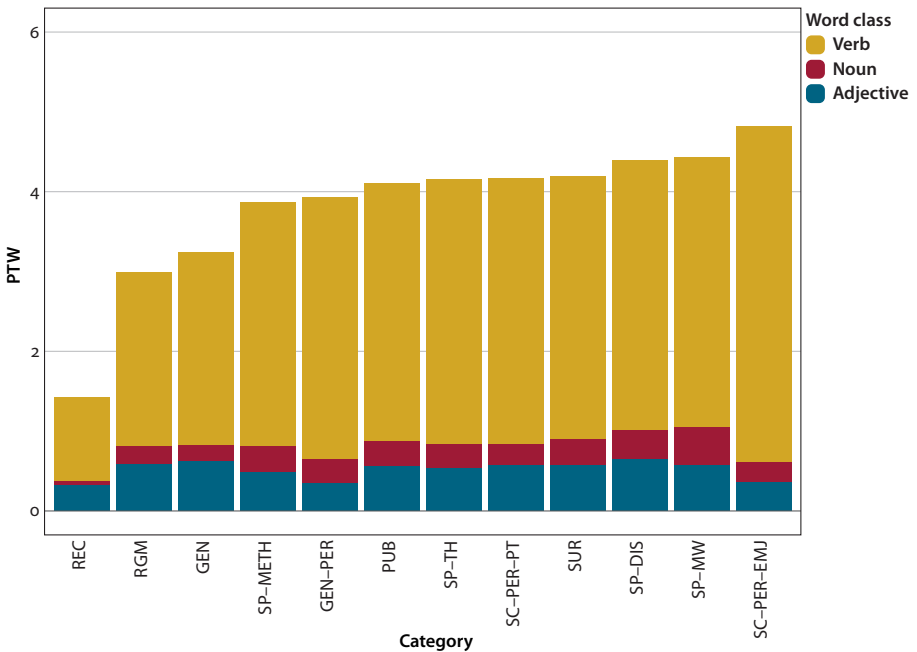
Figure 3. Types of stance *that*-clauses across categories (pooled frequencies)

11. Note that LMEMT comprises three different periodicals, whereas the findings of Gray, Biber, and Hiltunen (2011) is based solely on medical items published in the *Philosophical Transactions*, which are divided into three sub-categories.

Finally, we will look more closely at verbs, whose frequencies are plotted in Figure 4. This figure suggests that the relative frequencies of different verb types are similar across categories: around 90% of verb-licensed clauses are made up by three types of verbs (*certainty, likelihood,* and *communication*: *declare* verbs). The most frequently attested meaning type is *certainty*, which also matches Gray, Biber, and Hiltunen's (2011) finding for Early Modern English.
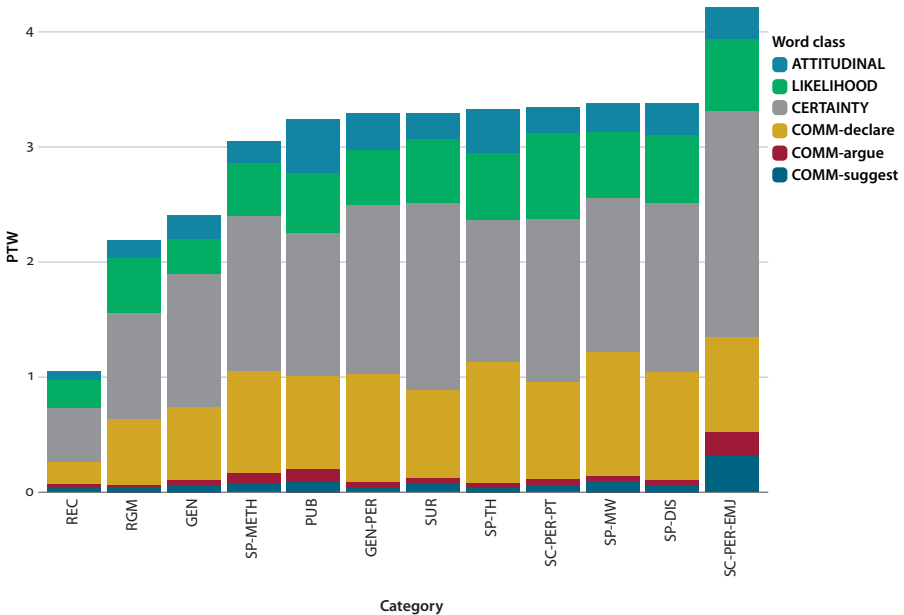
Figure 4. Types of licensing verbs (pooled frequencies)

We can further observe that *argue* and *suggest* verbs stand out as being proportionally more frequent in *Scientific Periodicals: EMJ* than in the other categories. This appears to be related to the dialogic nature of the journal articles, where statements and ideas of other authors are frequently relayed and assessed (Example 5). While detailed assessment of sub-register differences is beyond the scope of this study, it is nonetheless likely that the use of these verbs is also related to the genre of book reviews, which were observed to have a high frequency of communication verbs licensing *that*-clauses in Gray, Biber, and Hiltunen (2011).

(5)   He **alledges**, *that, in order to prepare every secreted liquor in the body, absorption from the secreting organ is absolutely necessary* […]   (SC-PER-EMJ, 1773)

To sum up, the analysis of frequency has established that the broad patterns of stance marking with *that*-clauses are similar to those in previous centuries. It thus

seems that despite the increasing specialisation and technologisation of eighteenth-century medicine, the style of argumentation still largely followed the patterns established in the previous centuries, as far as stance marking is concerned. This is particularly true for conservative text categories with a long history, such as *Medical recipe Collections* and *Regimens and health guides*, whereas some signs of convergence can also be observed between other categories.

The frequency-based analysis provides a bird's-eye view of stance marking, which, as previously argued, needs to be contextualised with the help of a more qualitative perspective. Similarly, while the analysis of verb semantics in Figure 4 enables the comparison with earlier results such as Gray, Biber, and Hiltunen (2011), it is not sensitive to differences between individual sentence meanings, as this requires a more elaborate analysis of clause subjects and other contextual features. To address these issues, we now turn to the analysis of local grammar.

5.   Local grammar analysis

5.1   Obtaining a sample

So as to extract a representative sample of verb-licensed stance *that*-clauses, the top verbs licensing *that*-clauses in the whole corpus are listed in Table 2. The table shows that the type-token ratio resembles a Zipfian distribution, in that the few very frequent verbs account for the vast majority of tokens, while there is also a large number of types (134) which only occur once. To extract an optimally representative sample, it is therefore reasonable to focus on the top verbs. The sampling frame includes thirteen verbs that occur more than 100 times in the dataset (*observe, say, find, tell, know, suppose, show, conclude, think, believe, inform, prove, see*), which together make up about 57% of all tokens of verb-licensed *that*-clauses (3,270 instances).

Table 2. Frequencies of verbs licensing *that*-clauses in LMEMT

	Verb	Frequency		Verb	Frequency
1	*observe*	608	8	*conclude*	187
2	*say*	417	9	*think*	179
3	*find*	348	10	*believe*	165
4	*tell*	287	11	*inform*	164
5	*know*	255	12	*prove*	151
6	*suppose*	205	13	*see*	111
7	*show*	194		(207 other verbs)	(2,513)
				Total	**5,783**

To obtain a representative sample of these tokens, 10% were randomly selected for closer analysis. This sample (328 tokens) was annotated for the three variables described above (SOURCE, TARGET, PROCESS TYPE), and this dataset forms the basis of the description of local grammar.

5.2   From meaning elements to discourse functions

The analysis of local grammar takes as its starting point the variable SOURCE, which is undoubtedly the most fundamental aspect of verb-licensed *that*-clauses, and importantly, one that is not captured by the frequency-based approach. By contrast, as this construction always evaluates some kind of a statement, idea or proposition, and not e.g. a person or a physical object, it is not surprising that the TARGET variable proves less distinctive. The third variable, PROCESS was already quantitatively analysed above (see Figure 4), and therefore it will only be commented on in the discussion of specific examples.

   We distinguish between statements based on what social actors they are attributed to, using the terms *averral* and *attribution* as per Sinclair (1986), Hunston (2000) and Charles (2006). At one end, we have **Emphasised averrals** (Example 6), where writers explicitly attribute the stance expression to themselves, typically using pronouns *I/we*). At the other end we find **Attributions** (7), where the evaluative act is explicitly attributed to someone else (e.g. using a third-person pronoun or a proper name). The remaining options are situated between these two ends: **Plain averrals** (8) make no explicit attribution, and **Hidden averrals** (9) attribute the statement to an inanimate entity:

(6)   […] when I had finished it, I **found** *that the cist was dependent from, or continuous with, a membranous tube* […]
<div align="center">(SC-PER-PT, 1764, Vol 54) (EMPHASISED AVERRAL)</div>

(7)   He certainly **thought** *that this calcination, like the incineration of vegetables, would produce an alkaline salt.*
<div align="right">(SP-DIS, Berkenhout, *An essay on the bite of a mad dog*, 1783) (ATTRIBUTION)</div>

(8)   It is well **known**, *that when emetics, neutral salts, and Opium are united, a much larger dose of the Opium may be administered with safety and success* […]
<div align="right">(SP-TH, Hamilton, *Practical Hints on Opium Considered as a Poison*, 1790) (PLAIN AVERRAL)</div>

(9)   These experiments **indicate**, *that breathing renders common air unfit for supporting life or flame, by depriving it of oxygene.*
<div align="right">(SP-METH, Beddoes, *Factitious Airs*, 1795) (HIDDEN AVERRAL)</div>

The frequencies and proportions of these types in the sample are shown in Table 3. As can be seen, the data represents a considerable variety of configurations. This already indicates the usefulness of detailed phraseological analysis as a complement to frequency-based study of stance, as the latter may not capture differences between individual lexicogrammatical patterns.

Table 3. Frequencies of source types of verb-licensed stance *that*-clauses in the sample

Type of source	Frequency	Percentage
ATTRIBUTION	126	38.4%
Doctor/medical author	60	18.3%
Patients/their companions	27	8.2%
Reader	18	5.5%
Reviewee	10	3.0%
Group	9	2.7%
Other	2	0.6%
PLAIN AVERRAL	62	18.6%
HIDDEN AVERRAL	29	8.8%
EMPHASISED AVERRAL	107	32.6%
Total	328	100%

Furthermore, it turns out that while stance *that*-clauses are commonly used for author-sourced claims, attributions are no less important. In fact, in terms of frequency, attributions are slightly more frequent than emphasised averrals. These two source types are the main SOURCE categories for stance *that*-clauses, accounting for over 70% of all instances. The remaining portion is split between the two minor types of averrals (plain averrals and hidden averrals) in a 2:1 ratio.

After identifying four broad categories of expressions which differ in terms of their SOURCE, we shall look more closely at their discourse-pragmatic functions and how they are realised lexicogrammatically. To this end, a main discourse function was identified for each sentence in the sample. The categorising process was iterative, starting with few basic categories and refining them over several rounds of analysis, with the aim of coming up with a classification that would be as simple as possible while still capturing the main differences. Following the conventions in some earlier work on local grammar (Hunston and Su 2017), the elements are summarised in Tables 4–6 and illustrated with examples taken from the data.

We shall begin with **attributions**, which can be further subdivided into a number of distinct groups of social actors. As seen in Table 3 above, these groups include other doctors and medical authors (both contemporaries and historical authors), patients and people in their immediate circle (e.g. family members, servants), (implied) readers of the text, and different groups of people. Given this variety, it is

not surprising that attributed stance *that*-clauses indeed serve a variety of discourse functions. The main functions attested in the data are summarised in Table 4; for each function, the associated social actor is also indicated.

Table 4. Main discourse functions for Attributions

Function	Associated social actor	Example(s)
Reporting statement by colleague	Medical author	(10), (11)
Relaying information about cases	Patient	(12)
Metadiscursive commentary	Reader	(13), (14)
Evaluating claim in medical book	Reviewee	(15)

The most characteristic function of attributions is to report and acknowledge a claim, statement or discovery by another medical professional, typically made in another medical text. This configuration favours communication verbs like *say* (10), but cognitive verbs are also frequently attested (e.g. *suppose* in Table 1 above). Example (11) shows how the verb phrase *endeavours to prove* can be used to report Dr. Johnstone's claim while at the same time attributing a specific intention to his expressing it in his book:

(10)   That eminent Physician Mr Lewis, **says**, *That of all Remedies, whether external or internal, there is none can equal the virtue of the Cold Bath* […]
     (GEN, Anonymous, *A Dissertation Upon The Nervous System*, 1780)

(11)   From many passages in the writings of Hippocrates, Celsus, and other ancients, Dr. Johnstone **endeavours to prove**, *that they were well acquainted with this disease.*   (SC-PER-EMJ, 1779)

Another prominent type is constituted by statements attributed to patients, which are clearly linked to medical case reports and are used by authors to explain how they obtained specific information relevant to a case. Typically these statements involve communication and cognitive verbs like *say, inform* and *observe*, and they deal with symptoms and effects of treatment, as in Example (12):

(12)   The patient only **observed**, *that it heated the fauces, and warmed the stomach, for a short time after it was taken.* (SC-PER-EMJ, 1789)

Attributions to the reader are clearly metadiscursive, in that writers use them to comment on how they expect their argument to be received. Grammatically, they are signalled either through the second-person pronoun *you* (13), or through third-person pronouns and nouns like *reader* (14). Modals like *will* and *may* are likewise frequently attested in these statements:

(13)   <u>You</u> may **observe** *that cool, fat, moist Bodies will bear hunger better than lean, dry, cholerick, hot Bodies, which perspire much*, […]
                    (REG, Byfield, *Directions tending to health and long life*, 1717)

(14)   I do not expect, that the reader should have any dependance on what I may have asserted in the preceeding pages; because, if he has any knowledge of mankind, <u>he</u> has reason to **suppose**, *that I have sacrificed truth to interest and a favourite hypothesis.*   (1767, SP-TH, Stern, *Medical Advice*)

Finally, Example (15) illustrates how attributions are used in book reviews to report and evaluate claims made by authors of medical books under review. These are similar to statements attributed to colleagues discussed above; the only difference is that they appear in Scientific Periodicals and are explicitly marked as book reviews. This functional similarity is further underlined by the fact that Example (15) specifically describes the reviewee's original intention, in the same way as in Example (11) above:

(15)   TO divest consumption of part of its horrors, and to **shew** *that it is curable in its worst stage*, is the professed intention of Dr. May in this little work.
                    (GEN-PER-GM, 1792)

Table 5. Main discourse functions for emphasised and plain averrals. The symbol ✓ indicates that the function is attested in the sample

Function	Emphasised Averral	Plain Averral	Example(s)
Making claim/reporting discovery	✓	✓	(16), (17)
Statement of current knowledge	✓	✓	(18)
Hypothetical reasoning	✓	✓	(19)

Next, we shall look at emphasised and plain averrals in more detail. These two types are fundamentally similar in that both express self-sourced claims, and therefore they are jointly summarised in Table 5. What distinguishes them is the degree of the writer's visibility and the strength of stance. The strongest stance is expressed through emphasised averrals with first-person singular pronoun *I* as the subject (Example 16), whereas the visibility is somewhat lower in statements with *we* as the subject, since there is a possibility for some ambiguity as to who is ultimately responsible for the claim (Charles 2006: 507–8). Thus in Example (17), the claim is carefully presented as an outcome of collective reasoning, based on due consideration of evidence. The fact that instances with *we* make up roughly a third of emphasised averrals in the sample testifies to the dialogicity of the process of knowledge construction in eighteenth-century medical writing. Even lower writer

visibility is found in plain averrals, which usually summarise the current state of knowledge rather than present original claims, as illustrated in (18). However, this is a tendency rather than a categorical distinction. The data also contains both plain averrals that make claims to new knowledge, and emphasised averrals summarising existing knowledge.

(16) For I **believe** *that it has scarce ever continued the same for a hundred years together.* (SP-TH, Heberden, *Antitheriaka*, 1745) (EMPHASISED AVERRAL)

(17) From which Data, I think, we **may safely conclude**, *that the Crassamentum, or Globular part of the Blood, consists of some Phlegm united with the Oil and Salts, and a small quantity of Earth.*
(SC-PER PT, 1717) (EMPHASISED AVERRAL)

(18) IT is well **known**, *that many important articles in the Materia Medica consist of metalic substances in a state of dissolution* […]
(GEN-PER-GM, 1761) (PLAIN AVERRAL)

Both types of averral frequently contain other markers of stance and evidentiality besides stance *that*-clauses, as already illustrated in (17), where the claim is presented as being based on data and hedged using the modal *may*, the adverbial *safely*, and the epistemic parenthetical *I think*. It can also be noted that not all instances are actual claims, but rather represent a process of hypothetical reasoning (Example 19):

(19) But I am afraid, if we come to examin the Matter a little more closely, we shall **find** *that this is generally all Stratagem*; […]
(SUR, Taylor, *An account of the mechanism of the eye*, 1727)

Finally, we shall briefly consider the use of hidden averrals, whose configurations are different from the other averral types. As indicated in Table 6, the two main types are *fact construction* and *paper construction*. These terms are borrowed from Dorgeloh and Wanner (2009) and refer to two types of an inanimate entity that occupy the subject position in a sentence: NPs encapsulating facts, observations, results, and the like (*fact construction*), and NPs focusing on the textual quality of the noun (*paper construction*). These are illustrated respectively in (20) and (21) below. Three verbs, *show, indicate* and *prove*, are attested in the sample:

Table 6. Main discourse functions for hidden averrals

Function	Example
Fact construction	(20)
Paper construction	(21)

(20)   These <u>experiments</u> **indicate**, *that breathing renders common air unfit for sup-porting life or flame, by depriving it of oxygene.*

<div align="right">(SP-METH, Beddoes, *Factitious Airs*, 1795)</div>

(21)   These <u>remarks</u> **may serve to show** *that the puerperal state does not strongly require the assistance of hospitals*; […]

<div align="right">(PUB, Aikin, *Thoughts On Hospitals*, 1771)</div>

The fact construction has been linked to the rise of object-oriented discourse (Atkinson 1999), and Dorgeloh and Wanner (2009) document a diachronic increase in its frequency in the *Science* section of the ARCHER (*A Representative Corpus of Historical English Registers*) corpus. By contrast, the paper construction is found to be a later innovation in scientific discourse, but the present analysis confirms that it was already used in eighteenth-century medical English.

To sum up, this data-driven description of local grammar of stance marking with verb-licensed *that*-clauses has clearly demonstrated that this grammatical construction occurs in a variety of semantic and pragmatic configurations in the corpus data, and that these configurations can be mapped onto specific textual meanings. Such textual meanings are identifiable due to their regular occurrence in the data, although the form-to-function correspondence is not one-to-one, and due to the multilayered nature of stance expression, individual categories inevitably comprise some degrees of variation. This is of course a common characteristic of all pragmatic phenomena. Importantly, as local grammar analysis systematically investigates aspects of stance expression that are not at the core of frequency-based approaches, it serves as a useful complement to previous corpus-based work by providing more nuance to the general patterns observed in corpora.

The compatibility of the two approaches, and the implications of the present findings to the analysis of stance in historical medical writing, will be discussed on a more general level in the final section of the chapter.

6.   Discussion and conclusion

This chapter has investigated the expression of stance with *that*-clauses using two approaches: a frequency-based approach following Biber et al. (1999) and an appli-cation of the notion of local grammar (e.g. Hunston and Sinclair 2000; Hunston and Su 2017). As previously suggested, these approaches are seen as complementary, and the present study thus integrates a quantitative, vertical reading methodology with a qualitatively-oriented, horizontal reading approach (cf. Rühlemann and Aijmer 2015). The former, frequency-based approach identifies broad lexicogram-matical patterns that are characteristic of the data and indicates at what rates they

occur. A more detailed appraisal of their textual and pragmatic functions is carried out using the local grammar approach, in which a systematically extracted sample of the data is subjected to close discourse-pragmatic and phraseological approach.

From the frequency analysis perspective, it can be observed that eighteenth-century medical writing represents a period of stylistic continuation rather than dramatic innovation, at least as far grammatical marking of stance is concerned. Frequency patterns are broadly similar to those of the two previous centuries, as described in previous studies. This provides an interesting contrast with the view of the eighteenth century as a transition period in medicine (e.g. Taavitsainen, Jones, and Hiltunen 2019), as well as the findings of Taavitsainen, Schneider, and Jones (2019), who identify several changes in the lexical make-up of the texts that correspond to contemporary ideological, societal, and disciplinary changes. Compared to these changes at the lexical level, it seems that changes in grammatical marking of stance are more muted, which of course is entirely reasonable. Lexical changes, especially those related to technological innovation and social change, are potentially quick and often also short-lived, whereas processes of grammatical change typically take a longer time to complete.

The main contribution of local grammar analysis is a more nuanced description of the discourse functions (particularly related to the SOURCE of statements), drawing attention to local textual meanings that may otherwise remain elusive. The variety of these meanings confirms the usefulness of this approach as a complement to frequency analysis. The relatively high incidence of attributions in the data, and the variety of social actors present in them, is clearly related to the dialogicity of Late Modern English medical argumentation and the importance of discourse communities, which has also been described in previous research (e.g. Atkinson 1999).

It should be noted that due to the broad scope of the chapter covering all categories of eighteenth-century medical writing as represented in LMEMT, the description of the local grammar elements in Section 5 is not exhaustive, and more detailed descriptions of the co-occurrence of stance *that*-clauses with other markers of stance and evidentiality as well as other grammatical and phraseological features like tense, modality and polarity would likewise be needed. For the same reason, the local grammars of adjective- and noun-licensed *that*-clauses were left out from the current analysis.

In general, what limits the applicability of local grammar analysis is the fact that the annotation is labour-intensive and therefore does not scale well to larger datasets. This issue is of course not limited to the analysis of stance, but applies more generally to the corpus-based analysis of pragmatic phenomena, and has been widely discussed in previous literature (e.g. Kohnen 2009; Jucker et al. 2008; Jucker and Taavitsainen 2014; Landert 2019). The same issue also limits the potential of

this approach for diachronic analysis, as this would require a substantial number of annotated concordance lines representing different points in time.[12] While the present analysis has mostly treated the LMEMT corpus as a synchronic sample, some indications of gradual historical change can nonetheless be detected in the data. This can be seen in Figure 5, which shows the percentages of the top thirteen stance verbs relative to the total frequency of verb-licensed *that*-clauses in four consecutive 25-year periods.[13] As can be seen, the largest relative increase can be attested for verbs of accomplishment, *find* and *observe*. This is consistent with the increasing importance of experimental philosophy and empiricism ever since the second part of the seventeenth century, as these verbs are typically used for reporting the results of experiments. On the other hand, communication verbs *tell* and *say* appear to be on the decline, which would also match the gradual waning of Aristotelian natural philosophy.
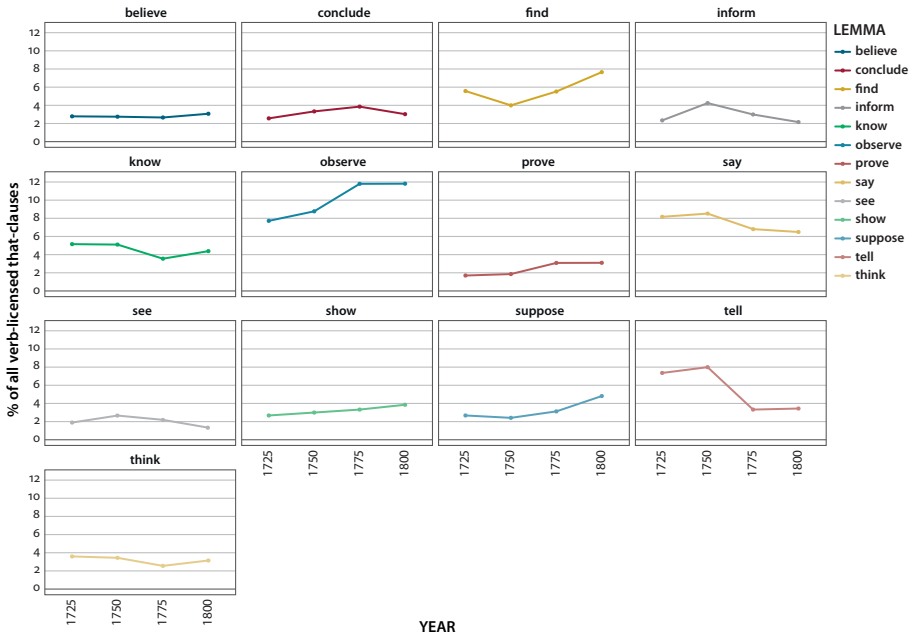
Figure 5. Percentages of top stance verbs (25-year bins)

---

12. Cf. Rissanen's (1989) "vanishing reliability" problem.

13. I.e. 1700–1725, 1726–1750, 1751–1775, and 1776–1800.

The existence of such frequency patterns may serve as a kind of "proof-of-concept" and provide additional support for the idea that stance-marking patterns can be meaningfully related to sociocultural and intellectual change in the field of medicine. However, as demonstrated throughout this chapter, functions of specific stance markers clearly vary considerably across different categories of writing as well as between individual texts. To account for this variation, it is necessary to apply a qualitative method which systematically addresses the contribution of different phraseological elements in conjunction with each other, and the present analysis has shown that local grammar analysis is a promising solution for this issue.

References

Primary sources

Taavitsainen, Irma, Turo Hiltunen, Anu Lehto, Ville Marttila, Päivi Pahta, Maura Ratia, Carla Suhr and Jukka Tyrkkö. 2019. *Late Modern English Medical Texts: The Corpus*. In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221

Secondary sources

Atkinson, Dwight. 1999. *Scientific Discourse in Sociohistorical Context: The Philosophical Transactions of the Royal Society of London, 1675–1975*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Barnbrook, Geoff. 2002. *Defining Language: A Local Grammar of Definition Sentences*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/scl.11

Bazerman, Charles. 1988. *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison, WI: University of Wisconsin Press.

Bednarek, Monika. 2008. *Emotion Talk Across Corpora*. Basingstoke: Palgrave Macmillan. https://doi.org/10.1057/9780230285712

Biber, Douglas. 2004. "Historical Patterns for the Grammatical Marking of Stance: A Cross-Register Comparison." *Journal of Historical Pragmatics* 5 (1): 107–36. https://doi.org/10.1075/jhp.5.1.06bib

Biber, Douglas, and Susan Conrad. 2019. *Register, Genre, and Style*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108686136

Biber, Douglas, and Edward Finegan. 1989. "Styles of Stance in English: Lexical and Grammatical Marking of Evidentiality and Affect." *Text* 9 (1): 93–124.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Charles, Maggie. 2006. "The Construction of Stance in Reporting Clauses: A Cross-Disciplinary Study of Theses." *Applied Linguistics* 27 (3): 492–518. https://doi.org/10.1093/applin/aml021

Charles, Maggie. 2007. "Argument or Evidence? Disciplinary Variation in the Use of the Noun *That* Pattern in Stance Construction." *English for Specific Purposes* 26 (2): 203–18. https://doi.org/10.1016/j.esp.2006.08.004

Dorgeloh, Heidrun, and Anja Wanner. 2009. "Formulaic Argumentation in Scientific Discourse." In *Formulaic Language, Volume II: Acquisition, Loss, Psychological Reality, and Functional Explanations*, ed. by Roberta Corrigan, Edith A. Moravcsik, Hamid Ouali, and Kathleen M. Wheatley, 523–44. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tsl.83.16dor

Finegan, Edward. 2009. "Legal Writing: Attitude and Emphasis. Corpus Linguistic Approaches to 'Legal Language': Adverbial Expression of Attitude and Emphasis in Supreme Court Opinions." In *The Routledge Handbook of Forensic Linguistics*, ed. by Malcolm Coulthard and Alison Johnson, 65–77. Abingdon: Routledge.

Garside, Roger, and Nick Smith. 1997. "A Hybrid Grammatical Tagger: CLAWS4." In *Corpus Annotation: Linguistic Information from Computer Text Corpora*, ed. by Roger Garside, Geoffrey Leech, and Anthony McEnery, 102–21. London: Longman. https://doi.org/10.4324/9781315841366-13

Goźdź-Roszkowski, Stanisław. 2017. "Signalling Sites of Contention in Judicial Discourse. An Exploratory Corpus-Based Analysis of Selected Stance Nouns in US Supreme Court Opinions and Poland's Constitutional Tribunal Judgments." *Comparative Legilinguistics* 32: 91–117. https://doi.org/10.14746/cl.2017.32.4

Goźdź-Roszkowski, Stanisław. 2019. "'It Is Not a Fact That the Law Requires This, but It Is a Reasonable Fact.' Using the Noun *That*-Pattern to Explore Stance Construction in Legal Writing." In *Corpus-Based Research on Variation in English Legal Discourse*, ed. by Teresa Fanego and Paula Rodríguez-Puente, 123–146. Studies in Corpus Linguistics. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/scl.91.06goz

Gray, Bethany, and Douglas Biber. 2012. "Current Conceptions of Stance." In *Stance and Voice in Written Academic Genres*, ed. by Ken Hyland and Carmen Sancho Guinda, 15–33. Hounsmill, Basingstoke: Palgrave MacMillan. https://doi.org/10.1057/9781137030825_2

Gray, Bethany Ekle, Douglas Biber, and Turo Hiltunen. 2011. "The Expression of Stance in Early (1665–1712) Publications of the *Philosophical Transactions* and Other Contemporary Medical Prose: Innovations in a Pioneering Discourse." In *Medical Writing in Early Modern English*, ed. by Irma Taavitsainen and Päivi Pahta, 221–47. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511921193.013

Harwood, Nigel. 2006. "(In)appropriate Personal Pronoun Use in Political Science: A Qualitative Study and a Proposed Heuristic for Future Research." *Written Communication* 23 (4): 424–50. https://doi.org/10.1177/0741088306293921

Hiltunen, Turo. 2010. Grammar and Disciplinary Culture: A Corpus-Based Study. PhD dissertation. Helsinki: Department of Modern Languages, University of Helsinki. http://urn.fi/URN:ISBN:978-952-10-6464-7

Hiltunen, Turo, and Irma Taavitsainen. 2019. "Towards New Knowledge: The Corpus of *Late Modern English Medical Texts*." In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Turo Hiltunen and Irma Taavitsainen, 1–16. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221.01hil

Hiltunen, Turo, and Jukka Tyrkkö. 2013. "Tagging *Early Modern English Medical Texts*." Paper presented at the *Corpus Analysis with Noise in the Signal (CANS)* workshop, Lancaster University, 22nd July 2013. http://ucrel.lancs.ac.uk/cans2013/

Hunston, Susan. 2000. "Evaluation and the Planes of Discourse: Status and Value in Persuasive Texts." In *Evaluation in Text: Authorial Stance and the Construction of Discourse*, ed. by Susan Hunston, and Geoff Thompson, 176–207. Oxford: Oxford University Press.

Hunston, Susan. 2003. "Frame, Phrase or Function: A Comparison of Frame Semantics and Local Grammars." In *Proceedings of the Corpus Linguistics 2003 Conference*, 342–58. Lancaster: UCREL.

Hunston, Susan, and John Sinclair. 2000. "A Local Grammar of Evaluation." In *Evaluation in Text: Authorial Stance and the Construction of Discourse*, ed. by Susan Hunston, and Geoff Thompson, 74–101. Oxford: Oxford University Press.

Hunston, Susan, and Hang Su. 2017. "Patterns, Constructions, and Local Grammar: A Case Study of 'Evaluation'." *Applied Linguistics* 40 (4): 567–93. https://doi.org/10.1093/applin/amx046

Hyland, Ken. 2005a. *Metadiscourse: Exploring Interaction in Writing*. London: Continuum.

Hyland, Ken. 2005b. "Stance and Engagement: A Model of Interaction in Academic Discourse." *Discourse Studies* 7 (2): 173–92. https://doi.org/10.1177/1461445605050365

Hyland, Ken, and Feng (Kevin) Jiang. 2016. "Change of Attitude? A Diachronic Study of Stance." *Written Communication* 33 (3): 251–74. https://doi.org/10.1177/0741088316650399

Hyland, Ken, and Carmen Sancho Guinda. 2012. "Introduction: A Context-Sensitive Approach to Stance and Voice." In *Stance and Voice in Written Academic Genres*, ed. by Ken Hyland and Carmen Sancho Guinda, 1–11. Hounsmill, Basingstoke: Palgrave MacMillan. https://doi.org/10.1057/9781137030825

Hyland, Ken, and Polly Tse. 2005a. "Evaluative *That* Constructions: Signalling Stance in Research Abstracts." *Functions of Language* 12 (1): 39–63. https://doi.org/10.1075/fol.12.1.03hyl

Hyland, Ken, and Polly Tse. 2005b. "Hooking the Reader: A Corpus Study of Evaluative *That* in Abstracts." *English for Specific Purposes* 24 (2): 123–39. https://doi.org/10.1016/j.esp.2004.02.002

Jaffe, Alexandra. 2009. "Introduction: The Sociolinguistics of Stance." In *Stance: Sociolinguistic Perspectives*, ed. by Alexandra Jaffe, 3–28. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195331646.001.0001

Jucker, Andreas H., Gerold Schneider, Irma Taavitsainen, and Barb Breustedt. 2008. "Fishing for Compliments: Precision and Recall in Corpus-Linguistic Compliment Research." In *Speech Acts in the History of English*, ed. by Andreas H. Jucker and Irma Taavitsainen, 273–94. John Benjamins Publishing Company. https://doi.org/10.1075/pbns.176.15juc

Jucker, Andreas H., and Irma Taavitsainen. 2014. "Diachronic Corpus Pragmatics: Intersections and Interactions." In *Diachronic Corpus Pragmatics*, ed. by Irma Taavitsainen, Andreas H. Jucker, and Jukka Tuominen, 3–26. John Benjamins Publishing Company. https://doi.org/10.1075/pbns.243.03juc

Kohnen, Thomas. 2009. "Historical Corpus Pragmatics: Focus on Speech Acts and Texts." In *Corpora: Pragmatics and Discourse. Papers from the 29th International Conference on English Language Research on Computerized Corpora (ICAME29)*, ed. by Andreas H. Jucker, Daniel Schreier, and Marianne Hundt, 13–36. Amsterdam: Rodopi.

Kuo, Chih-Hua. 1999. "The Use of Personal Pronouns: Role Relationships in Scientific Journal Articles." *English for Specific Purposes* 18 (2): 121–38. https://doi.org/10.1016/S0889-4906(97)00058-6

Landert, Daniela. 2019. "Function-to-Form Mapping in Corpora: Historical Corpus Pragmatics and the Study of Stance Expressions." In *From Data to Evidence in English Language Research*, ed. by Carla Suhr, Terttu Nevalainen, and Irma Taavitsainen, 169–90. Language and Computers. Leiden: Brill.

McEnery, Tony, and Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

McGrath, Lisa, and Maria Kuteeva. 2012. "Stance and Engagement in Pure Mathematics Research Articles: Linking Discourse Features to Disciplinary Practices." *English for Specific Purposes* 31 (3): 161–73. https://doi.org/10.1016/j.esp.2011.11.002

Rissanen, Matti. 1989. "Three Problems Connected with the Use of Diachronic Corpora." *ICAME Journal* 42 (1): 9–12. https://doi.org/10.1515/icame-2018-0002

Rühlemann, Christoph, and Karin Aijmer. 2015. "Corpus Pragmatics: Laying the Foundations." In *Corpus Pragmatics: A Handbook*, ed. by Karin Aijmer and Christoph Rühlemann, 1–26. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139057493.001

Sinclair, John. 1986. "Fictional Worlds." In *Talking About Text: Studies Presented to David Brazil on His Retirement*, ed. by Malcolm Coulthard, 43–60. Discourse Analysis Monograph. Birmingham: English Language Research.

Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Su, Hang. 2018. "'Thank Bloody God It's Friday': A Local Grammar of Thanking." *Corpus Pragmatics* 2 (1): 83–105. https://doi.org/10.1007/s41701-017-0024-9

Su, Hang, and Naixing Wei. 2018. "'I'm Really Sorry About What I Said': A Local Grammar of Apology." *Pragmatics* 28 (3): 439–62. https://doi.org/10.1075/prag.17005.su

Taavitsainen, Irma, and Turo Hiltunen (eds). 2019. *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.221

Taavitsainen, Irma, Turo Hiltunen, Anu Lehto, Ville Marttila, Päivi Pahta, Maura Ratia, Carla Suhr, and Jukka Tyrkkö. 2019. "*Late Modern English Medical Texts*: The Corpus". In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221.02taa

Taavitsainen, Irma, Peter Murray Jones, and Turo Hiltunen. 2019. "Sociohistorical and Cultural Context of *Late Modern English Medical Texts*." In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen, 17–30. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221.02taa

Taavitsainen, Irma, Gerold Schneider, and Peter Murray Jones. 2019. "Topics of Eighteenth-Century Medical Writing with Triangulation of Methods: LMEMT and the Underlying Reality." In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen and Turo Hiltunen, 31–74. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221.03taa

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/scl.6

"Die Blumenzeit der Frau"

A corpus-based study of the development of medical references to menstruation in historical texts on herbology

Gohar Schnelle, Carolin Odebrecht, Anke Lüdeling,
Laura Perlitz and Catharina Fisher
Humboldt University of Berlin

Scientific herbal texts in vernacular German first emerged in the 15th century, and their diachronic analysis presents an opportunity to trace linguistic changes across centuries. This study deals with the linguistic strategies that are used to refer to the concept of menstruation, which we study in the RIDGES corpus, representing herbal texts from the late fifteenth century to the early twentieth century. This exploratory study focuses on terminology, i.e. expressions referring to concepts relevant to scientific communication, and investigates the morphological, syntactic and semiotic characteristics of references to menstruation. Our study reveals two tendencies: a decrease of formal variation as well as a decrease of semiotic variation. Both tendencies point towards the increase of terminologisation in the early and late modern period (1483–1914).

Keywords: variation, corpus linguistics, technical terms, terminology, language change, lexicon

1. Introduction

The broader context of the present study is the emergence and change of the linguistic properties of scientific herbal texts in vernacular German during the early and late modern period (1483–1914), and our specific focus is on the linguistic strategies used to refer to medical concepts, specifically the concept of *menstruation*. The data come from the *Register in Diachronic German Science Corpus* (RIDGES, Odebrecht et al. 2017), a corpus of herbal texts from the late fifteenth century to the early twentieth century (with our focus on the early modern period). There is a great deal of variation in the use of terminology in the beginning of the period

in focus, with expressions such as *Heimlichkeit* ('secrecy') or the metaphorical *Blumenzeit* ('flower time') being used. Towards the end of the nineteenth century, the usage converges towards *Monatszeit* ('monthly time'). These linguistic changes reflect both the emergence of the scientific register and extralinguistic developments. Education becomes more widely available to larger parts of society, scientific concepts and methods evolve, and the status of the vernacular together with the availability of grammars, lexicons, or text books improves. All this contributes to changes in the relationship between the writer (expert) and the audience (experts and laypersons).

Our study focuses on the nominal reference to the physiological phenomenon of menstruation, and interprets linguistic changes against changes in society, such as the development of the scientific society as a social system (Zeige 2011). Using the RIDGES corpus, which is a diachronic historical corpus consisting of text sections of different length with the common topic of "herbology",[1] we investigate the process of terminologisation of references to menstruation by asking the following questions:

RQ 1:   What variants of referring expressions to menstruation are found in the corpus, and in what frequencies do they occur?

RQ 2:   In which ways is the existing vernacular or classical lexical material applied to create new referring expressions?

RQ 3:   How are the lexical, semiotic, morphological and syntactic variants distributed in our corpus and how does this distribution develop over time?

RQ 5:   Which extra-linguistic factors have an impact on the choice of variants and naming strategies?

We will also discuss the broad methodological question of how issues of corpus design and sampling influence the interpretation of findings.

We present an overview of the variation of referring expressions to menstruation on morphological, syntactic and semiotic levels. These relate to extra-linguistic historical developments. We also discuss possibilities to overcome restrictions caused by the corpus design.

---

1.   RIDGES (*Register in Diachronic German Sciences*) has been created in the course of different projects and in Master's and Bachelor's programs in the Faculty of Linguistics, Literature and Humanities at Humboldt-Universität zu Berlin since 2011. We would like to thank all students who have contributed to this resource. See Section 2 for more information on the corpus data.

2.   Referring expressions to menstruation in historical medical writing

Menstruation is a recurrent phenomenon for women. Even though the exact biological mechanisms were not well understood in early modern times, regular menstruation was recognised as a bodily process crucial for a woman's health and fertility (Kruse 1999; Evans 2012). Due to many problems ranging from pain or heavy bleeding to amenorrhoea, references to menstruation are frequent in herbal and medical texts.[2] Within the framework of Galenic medicine, menstruation was seen as a symptom of excess of blood in a woman's body. This excess was seen as a consequence of a woman's inactive lifestyle, as material for the genesis of the embryo or as nutrition for the embryo (King 2015). These ideas are seen in the concept of *corrupted menstruation.* As late as 1739, we find the lexicon entry *Monats-Reinigung, gantz verdorbene* ('monthly cleansing, entirely corrupted') (Zedler 1739),[3] reflecting the idea that a missed or delayed period is healed by the body by releasing blood from other body parts (e.g. nose, mouth, knee, legs). The herbal and medical texts in focus here mainly deal with problematic or pathological aspects of menstruation, such as its absence (e.g. *verstopfte Monats-Zeit* 'congested monthly time') or heavy menstrual bleeding (e.g. *übermåſzige Monat-Zeit* 'excessive monthly time'). We do not differentiate between medical contexts, but focus on the variation and linguistic referring expressions to the medical concept of *recurrent non-pathotological bleeding from the uterus.*[4] The medical need for describing cures for menstrual problems went hand in hand with socio-cultural aspects as demonstrated by indirect metaphorical references, such as *Blume* ('flower'), *Blumenzeit* ('time of flower'), *Heimlichkeit* ('secrecy') and *Blödigkeit* ('weakness').

The RIDGES corpus, which is the source of data for this chapter, only contains printed texts. The invention of movable type printing in the mid-fifteenth century brought about a media revolution, and the rapidly increasing production of texts

---

2.   Menstruation is related to all aspects of female sexuality (sexual desire, wanted and unwanted pregnancy, abortion, loss, etc.). Since our focus is linguistic, we will only mention aspects that are important to the understanding of certain expressions. See Evans (2012) for a detailed account of the history of menstruation.

3.   For more on Zedler's lexicon, see Section 4.1.

4.   Since we do not assume that the authors of our texts follow a scientific-medical approach in the conception of the phenomena described, we explicitly do not use a medical definition as a basis. A genuine medical definition, e.g. "mit Blutung einhergehende Abstoßung des Stratum functionale der Gebärmutterschleimhaut (am Anfang eines jeden Menstruationszyklus) als Hormonentzugsblutung nach ovulator. biphasischem Zyklus" (Tiemann et al. 2020: 1310–1311) implies precise knowledge about the underlying physiological causes and processes, which cannot be assumed in our texts. Instead, we define menstruation based on observable symptoms, since we assume that the authors also conceived the medical phenomenon on the basis of observations.

gave rise to a new vernacular readership interested in medical topics. Vernacular writers could then also pass on the knowledge of healers, wound doctors, bathers, midwives and herbalists to a supra-regional vernacular audience. As in many other countries in (Western) Europe, this led to the gradual vernacularisation of medical writing.[5]

The slow shift from the Latin academic tradition to the vernacular and the interweaving of academic and practical traditions of medical knowledge posed challenges on all linguistic levels (Klein 2011a). Writers had to find or invent vernacular terms that would be understood across different dialects. Linguistically, the attempt to find equivalent lexemes for Latin terms and to ensure supra-regional comprehensibility is reflected in a seemingly redundant range of extensional synonyms. Many of the texts printed in the fifteenth and sixteenth century were produced explicitly with the aim of combining these divergent traditions of knowledge (Caflisch-Schnetzler 2018; Katschning 2018). This intention was expressed in a number of prefaces, e.g. in the *New Kreuterbuch* by Leonhard Fuchs (1543). Fuchs criticises the apparently common fact that physicians could hardly distinguish between ten different herbs and devalued the knowledge of herbalism. Herbology, with its everyday relevance, was one of the first branches of science dealt with in the vernacular languages (cf. Riecke 2004; Gloning 2007; Taavitsainen and Pahta 2010, among many others). While early herbal books written in the vernacular mainly address concrete practical needs, later herbal books can be put into the context of Enlightenment and its demands of comprehensibility of scientific content (Klein 2011a; Lindner 2018; Roelcke 2018). The early herbals (often collations and translations of Latin and Greek medical and herbal texts) are precursors of texts in the emergent fields of botany, pharmacy, chemistry, evidence-based medicine as well as the popular science herbals we can still find today.

This change in the social function of the vernacular led to metalinguistic discussions in the seventeenth and eighteenth century (see Klein 2011b for an overview), e.g. about the appropriateness of using the vernacular in scientific texts and the requirements for an ideal scientific language. The focus of this chapter is on terminology, that is, the expressions which refer to concepts that are relevant in scientific communication; terminology is thought to be crucial for scientific writing (Gläser 1995–1996). *Terms* differ from non-terminological expressions in (a) their deliberate creation at a certain moment in time and (b) regarding their semiotic properties, i.e. the relation between their form and the concept being referred to (Reiner 2003). The creation of terminology is understood to be a deliberate

---

5. Note that this change began at slightly different times for different European speech communities (Guentherodt 1986) and was a slow process. It took about 300 years for scientific texts (Pörksen 1998), and it did not affect all registers at the same speed (Klein 2011a).

normative act (by an individual or a community) which follows a process of variant reduction, variant selection and standardisation.

3. Corpus and annotation

Our corpus study is based on the RIDGES Herbology Corpus Version 9.0 (Odebrecht et al. 2020).[6] It is a diachronic multi-layer corpus of historical German printed texts on herbs, medical preparations, and diseases, and it contains 305,056 *dipl* units (historical word forms; see below).

The corpus is deeply annotated on many linguistic levels. Table 1 illustrates the annotation layers.

Table 1. Annotation example in RIDGES. The extract comes from *Paradeiszgaertlein* (Rosbach 1588).[7]

dipl	Den	Frawen	bringts		jr	Blume	zeit
norm	Den	Frauen	bringt	es	ihr	Blumenzeit	
lemma	die	Frau	bringen	es	ihr	Blumenzeit	

Table 1 shows the annotation extract of the sentence *Den Frauen bringt es ihr Blumenzeit*, lit. 'to the women it brings their flower time'. The corpus contains three text representations: a diplomatic transcription (*dipl*) and two normalisations (*clean* and *norm*). The present study uses the *norm* layer, which contains manually normalised word forms (see Odebrecht et al. 2020 for detailed annotation guidelines). The corpus architecture allows the mapping of different tokenisations: The cliticised segment *bringts* ('it brings'), spelled as one word/unit in the original document, is diplomatically transcribed as one token while being represented by two tokens in *norm* (*bringt es*). *Blume zeit* ('flower time'), represented as two units on the *dipl* layer, is represented by the compound *Blumenzeit* on the *norm* layer.[8] For each *dipl* unit, the language is manually annotated.

---

6. Corpus version 9 is available through the ANNIS search and visualisation interface (Krause 2019). The corpus reference link is: https://korpling.org/annis3/#_c=UklER0VTX0hlcmJvbG9ne V9WZXJzaW9uOS4w.

7. *Dipl* covers the diplomatic transcription of the print, *clean* is an automatic normalisation based on characters, *norm* means manual phonological and morphological normalisation, and *lang* is the annotation of the language for each *dipl* unit. Annotation visualisation was done with ANNIS (Krause and Zeldes 2016).

8. The corpus representation thus has three minimal tokenisation layers, i.e. multiple segmentations, cf. Odebrecht et al. (2017).

3.1   Annotating referring expressions to menstruation

Annotating references to menstruation in a historical corpus is challenging. There are no established lists of possible terms and many of the possible words are ambiguous and require disambiguation. Therefore, manual annotation that takes the linguistic context into account is required. Example (1) illustrates the problem with the word *Zeit* ('time'):

(1)   a.   *Im Herbst <u>zur Zeit der Weinlese</u> macht man Wermutwein*
              'In autumn <u>at the time of the grape harvest</u>, we make Wormwood wine'
                                          (Friedrich Losch, *Kraeuterbuch*, 1914)
      b.   *auch säubert es den Frauen <u>ihre Flüsse</u> / und Stabwurzel mit Wein getrunken / bringt ihnen <u>ihre Zeit</u>*
              'It cleans <u>the flows of the women</u> and if they drink southernwood with wine it brings <u>them their time</u>'
                                          (Johannes Tallat, *Artzney Buchlein Der Kreutter*, 1532)
      c.   *Item Majoranwasser getrunken / fördert den <u>Weibern ihre Zeit</u>*
              'So if they drink majoram water it brings <u>the women their time</u>'
                                          (Adrian Nasser, *Thesaurus Sanitatis*, 1673)

In Example (1a) *Zeit* with a definite article refers to a period in time when a certain event ('grape harvest') takes place. It is thus clearly not what we look for. Example (1b) shows that the same lexeme with the possessive pronoun *ihr* ('her') refers to menstruation. Here, we analyse the anaphoric relations and the content as indicators for identifying references. Furthermore, we consider the tendencies of co-occurrence. These tendencies are diagnostically based on examples such as (1c). We can conclude that menstruation is meant with *Zeit* ('time') because the verb *bringen* ('bring') selects the indirect object *den Weibern* ('the women') and direct object *ihre Zeit* ('their time').

The references to menstruation were manually annotated by four trained and experienced annotators, following to the annotation schema described below. Each extract was annotated by at least two annotators. The schema focuses on annotating each instance of a noun phrase that refers to menstruation and classifies its grammatical units and its status as a term, as shown in Table 2. The schema has separate annotation layers for morphology and syntax (layer *gram*) as well as semiotic information (layer *technicalterm*).

Table 2. Annotation example for a reference to menstruation in spreadsheet format from *Paradeiszgaertlein* (Rosbach 1588).[9]

dipl	clean	norm	M	gram	technicalterm
Den	Den	Den			
Frawen	Frawen	Frauen			
		bringt			
bringts	bringts	es			
jr	jr	ihr		det	
Blume	Blume				
zeit	zeit	Blumenzeit	M	nkc	met

Table 3 shows the annotation values of *gram*.[10] The noun kernel is marked according to its morphological complexity and the existence and category of its modifier(s).[11]

Table 3. Tag set for the annotation of *gram*

Tag	Description	Example
Nks	noun kernel simplex	*Zeit* ('time')
Nkd	noun kernel derivation	*Heimlichkeit* ('secrecy')
Nkc	noun kernel compound	*Blumenzeit* ('time of flowers')
attAdj	attributive adjective	*verstopfte Blume* ('congested flower')
genAtt	genitive attribute	*der Frauen Zeit* ('the women's time')
relAtt	relative attribute	*Blume die zu sehr fließt* ('flower that flows too much')

The annotation of the semiotic relation between the phrase and the phenomenon are marked in the *technicalterm* layer, following Pörksen (1998), who describes semiotic levels of abstraction between linguistic references and scientific concepts. The levels are on a scale between everyday language and scientific language. Table 4

---

9. The span annotation *M* for menstruation covers the minimal noun phrase (excluding syntactically higher embedded phrases). Annotation *gram* assigns grammatical units within the nominal phrase. It further classifies the noun kernel according to its morphological complexity. Each noun kernel and its modification is annotated in the layer *technicalterm* with selected term categories according to Pörksen's schema (1998).

10. We cannot discuss determiners in terms of definiteness here. Definiteness is a complex linguistic phenomenon and would need more elaborated grammatical description in the broader textual context.

11. We also integrated the category 'prepositional attributes' into our annotation guidelines. Since we did not find any instances in our corpus, we excluded this category from our discussion.

shows the three basic categories: paraphrases, metaphors and loan words.[12] The minimum level of abstraction (paraphrases) is assigned to the pole of everyday language, and the maximum level of abstraction (symbols) to the pole of scientific language.

Table 4. Tag set for the annotation of *technicalterm*

Tag	Description	Example
para	paraphrase	*Blutfluss* ('blood flow')
met	metaphor	*Blume* ('flower')
loanmed	medical loan word	*menstruum* ('menstruation')
her	inherited vocabulary modifying the reference	*der Frauen* ('women's')

In order to distinguish between paraphrases and metaphors, we considered whether obvious aspects of menstruation could be identified. Example (2a) paraphrases observable aspects of the phenomenon, as blood flowing is a strong indicator for menstruation. Similar aspects are interval, cycle, wetness, femininity, redundancy, genitalia, and illness. Example (2b) uses another semantic domain (*flower-botany*) to refer to aspects that are culturally associated (blossom) or indirectly refer to menstruation (being fertile, colour of the blood).

(2) a. *Blutfluss* ('blood flow'): flowing of blood => paraphrase
    b. *Blume* ('flower-botany'): fertility, colour, blossom => metaphor

3.2   The impact of corpus design

Ideally, corpus design should be informed by research questions or research context (cf. Hunston 2009; Lüdeling et al. 2011).[13] The original research question behind the RIDGES corpus is the following: how does German as a scientific language evolve and change after printing was invented? This question defines three selection criteria, namely language (German), medium (print), and period (1450–1920). The fourth criterion (topic) is less obvious. We decided on herbal texts because they are available from early on and in large numbers.

---

12. At this point, we do not understand the development of terminologisation well enough to use the full spectrum of Pörksen's (1998) schema.

13. For a broader discussion on corpus design and limitations of statistical representativeness, see Biber (1994) and Koplenig (2019) and for literary concepts on representativeness, see Bode (2018) as well as Algee-Hewitt and McGurl (2018).

While our sample is fairly large, we cannot say anything about its representativeness, as we did not conduct an extensive preliminary study to examine prints, publication places and publication history in this period and of this topic. RIDGES is therefore best described as an opportunistic corpus.

RIDGES contains 73 German text extracts, which come from 41 books. This means that some books are represented by several extracts.[14] In order to monitor the thematic range, RIDGES contains metadata about the general topic of the book. The majority of the extracts are classified as botanic and medical topics. Figure 1 shows the proportion of text material (normalised word forms) per year.
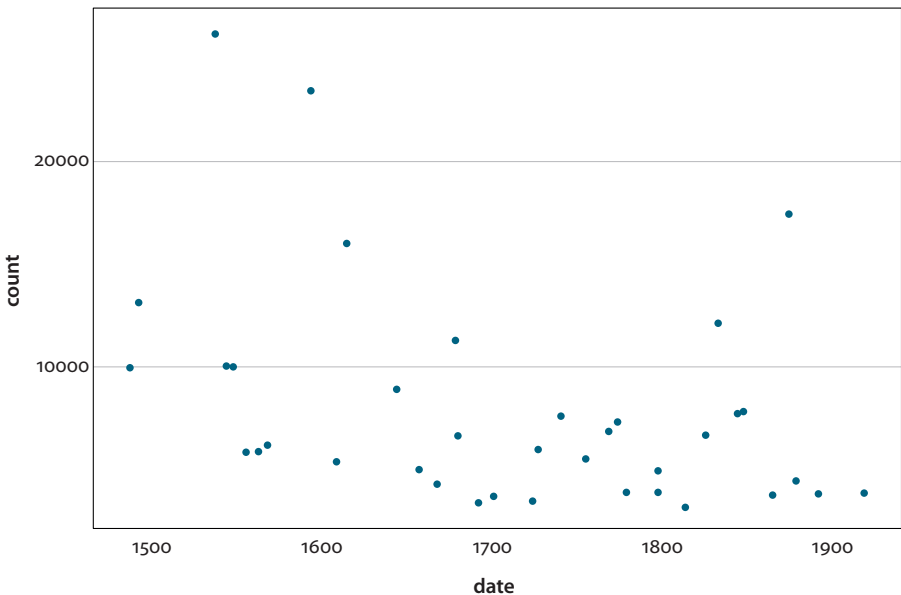
Figure 1. Distribution of *norm* units per year in the corpus. The dots represent the number of *norm* units per year

---

14. Because of the large amount of manual work that went into the transcription and annotation of the corpus, we were not able to cover everything we intended. At the start of the project in 2011, tools and services for automatic OCR for broken scripts were not available (for recent developments of OCR on blackletter prints, see Reul et al. 2019 and Springmann and Lüdeling 2017). We decided to include a higher number of shorter extracts from different books from different times rather than a few entire books.

Figure 1 shows that the corpus has an imbalance between the number of texts and the years of publication. The peaks indicate that there are longer or more extracts from a single book than from different books.[15] The chronological distribution is not balanced as not all years or decades have the same amount of *norm* units and not every year has a text.

Due to its opportunistic corpus design and its lack of balance between time, space, and communicative function, RIDGES is not suitable for a statistically valid diachronic study on variant distributions and factors. Despite these caveats, it can be used to obtain an exploratory overview of occurrences, variance, linguistic structures and extra-linguistic contexts of pre-scientific references to menstruation.

4. Corpus study

Our corpus study is based on the normalised word forms (*norm*), the annotation of references to menstruation (see Section 3.1) and the corpus metadata (*date*).

4.1 Distribution and variants of references to menstruation

RIDGES contains 196 noun phrases in total that refer to menstruation, spread across 34 documents.[16] The distribution of counts per year is shown in Figure 2.[17]

Figure 2 shows that the references to menstruation (196) are distributed over the entire time period of 400 years and that in 50% of the extracts we find fewer than five occurrences. The corpus design (see Section 2.2) must be taken into account here. The peaks result from the occurrences in *Contrafayt Kreuterbuch* by Otto Brunfels (1532) and *Thesaurus Sanitatis* by Adrian Nasser (1673), showing that a single document can have a strong influence on the frequency of referring expressions to menstruation. Moreover, the occurrences are not distributed equally within the documents. For example, only one of the two extracts from *Thesaurus Sanitatis* contains references to menstruation, whereas the occurrences in *Contrafayt Kreuterbuch* are distributed across all extracts. This illustrates how the distribution is in some way dependent on chapter content. For example, it is

---

15. The following books are represented by several extracts: *Contrafayt Kreuterbuch* (Brunfels, 1532, Straßburg), *Artzney Buchlein Der Kreutter* (Tallat, 1532, Leipzig), *Gart Der Gesundheit* (von Cuba, 1487, Ulm), *Sonderbares Kraeuterbuch* (Anonymous, 1675). The extract from *Deutsche Pflanzennamen* (Grassmann, 1870, Stettin) is relatively long compared to other extracts.

16. Because of different chapter lengths and page layouts in early prints, normalisation according to text length is not suitable.

17. Note that a particular year might contain several extracts.
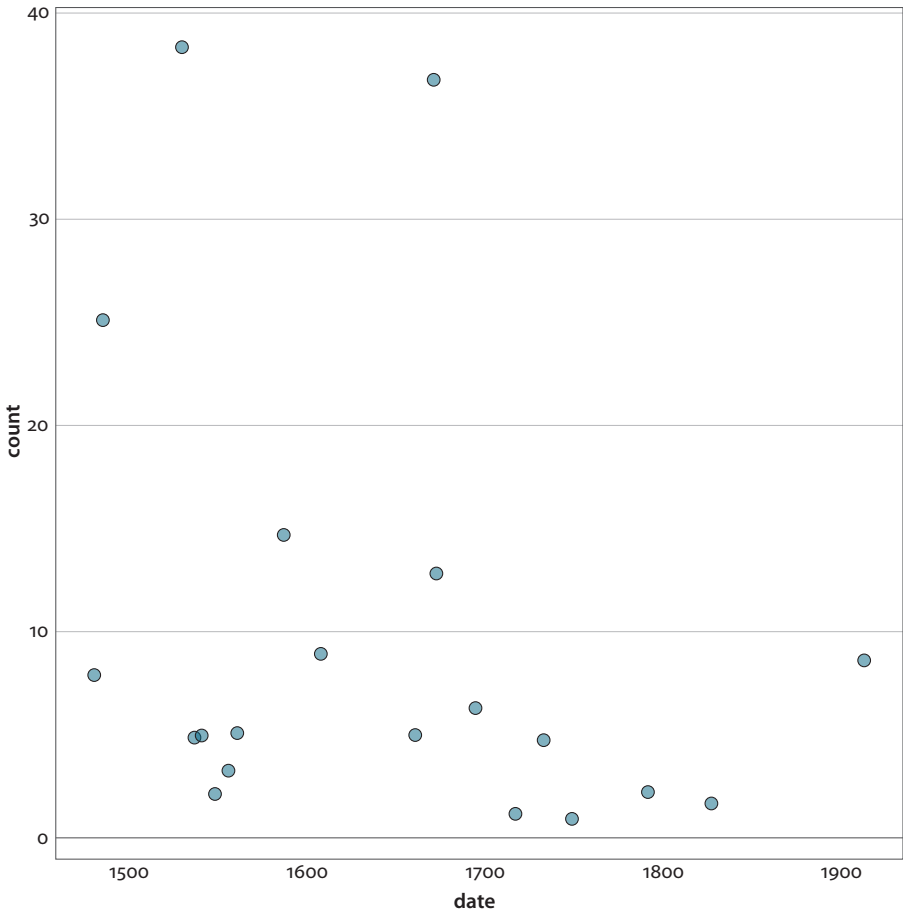
Figure 2. Distribution of noun phrases referring to menstruation

possible that some herbs are used to treat respiratory diseases and offer no relief for menstrual problems.

Figure 2 also shows a gap between the middle of the eighteenth century and the end of the nineteenth century. In this period only three references to menstruation can be found. This may simply be due to the fact that fewer texts from these years have been included in the corpus.[18] On the other hand, several references to menstruation are found in the first half of the sixteenth century, because two books from 1532 (*Artzney Buchlein der Kreutter* by Johannes Tallat and Brunfels' *Contrafayt Kreuterbuch*) are represented by three and four text extracts, respectively.

---

18. The corpus is being extended to include more texts from these periods.

The fact that almost 200 references to menstruation can be attested in the corpus also makes it possible to investigate the variants and types of references. Table 5 includes the 15 most frequent references to menstruation in the whole corpus.[19]

Table 5. Frequencies of norm units referring to menstruation

Norm-unit	Count
Zeit ('time')	49
Monatszeit ('monthly time')	18
Fluss ('flow')	14
Menstruum ('monthly')	12
Sucht ('disease')	9
Blumen ('flowers')	9
Blume ('flower')	6
Heimlichkeit ('secrecy')	6
menstruum ('monthly')	6
Blödigkeit ('weakness')	5
Krankheit ('disease')	5
Menses ('months')	4
Reinigung ('cleansing')	4
Menstrua ('monthly sacrifice')	3
Recht ('right')	3

The table shows that there is a great deal of terminological variation. The most frequent nouns referring to menstruation in the corpus are *Zeit* ('time'), *Monatszeit* ('monthly time') and *Fluss* ('flow'). In addition, metaphors such as *Blume* ('flower') or Heimlichkeit ('secrecy') are also frequent. These metaphorical expressions are not commonly used in modern German, but loan words such as *menstruum* or *menses* are among the most frequent word forms.

Table 6 lists the distribution of variants in two texts, which are separated by about 240 years.

As can be seen, the variety of terms is much greater in the earlier text, *Thesaurus Sanitatis,* containing 27 references and 13 variants. The later text, Losch's *Kräuterbuch*, contains 9 references and only 3 variants. The fact that we still find 3 variants in *Kräuterbuch*, which comes from a period in which scientific conventions were already established and terminologisation had taken place (Lindner 2018), is somewhat surprising. However, *Kräuterbuch* has a botanical focus and physiological phenomena may have been of marginal interest. Varying terminology for peripheral concepts is a common characteristic of scientific language even in modern texts (Gardt 1998).

---

19. As many of the words are ambiguous, we give only the most salient translation.

Table 6. The references to menstruation in two extracts

Nasser: *Thesaurus Sanitatis* (1673)		Losch: *Kräuterbuch* (1914)	
Zeit ('time')	18	*Monatszeit* (time of the month)	6
Fluss ('flow')	4	*Zeit* ('time')	2
Recht ('right')	3	*Reinigung* ('cleansing')	1
Monatszeit ('monthly time')	2		
Blume(n) ('flower')	2		
Dünste ('vapor')	1		
Weiberblumen ('women's flowers')	1		
Überfluss ('superfluity')	1		
Lässe ('phlebotomy')	1		
Blutfluss ('blood flow')	1		
Gerechtigkeit ('justice')	1		
Weiberfluss ('women's flow')	1		
Bewegungen ('movements')	1		

To assess the process of terminologisation further, we need to look for evidence for a terminological status of the references within the period. Are any of these expressions used in contemporary methodological discourse? Are any of them included in a reference framework, which would be the first indicator of standardisation or canonisation? These questions cannot be answered satisfactorily by this exploratory study, but some evidence can be gleaned from *Universallexikon* by Johann Heinrich Zedler (1731–1754). It is one of the most comprehensive encyclopaedic projects of the German Enlightenment (Caflisch-Schnetzler 2018). *Universallexikon* lists many of the lexical types in Table 5 as lemmata, but only with a pointer to the lemma *menstruum*, under which the detailed article is found. The Latin technical term *menstruum* (18 occurrences with two different spellings in RIDGES) is fourth on the frequency list. However, no evidence for the modern form *Menstruation* is found in the entire corpus. The fact that the lemma is not listed in the *Universallexikon* supports the assumption that at least until the release of the *Universallexikon*, it was not a common term.

To sum up, we have a variety of references to menstruation across the entire period of 400 years. The earliest texts demonstrate a high degree of variability, which seems to be decreasing over time. In addition, the first occurrences of new technical terms can be attested, which can be contextualised with the help of historical evidence.

4.2 Grammatical strategies to refer to menstruation

Next, we consider the distribution of the first potential candidates for technical terms and the structure of references. We start by looking at simple and complex words with the frequent roots *zeit* ('time') and *fluss* ('flow') in Figure 3.
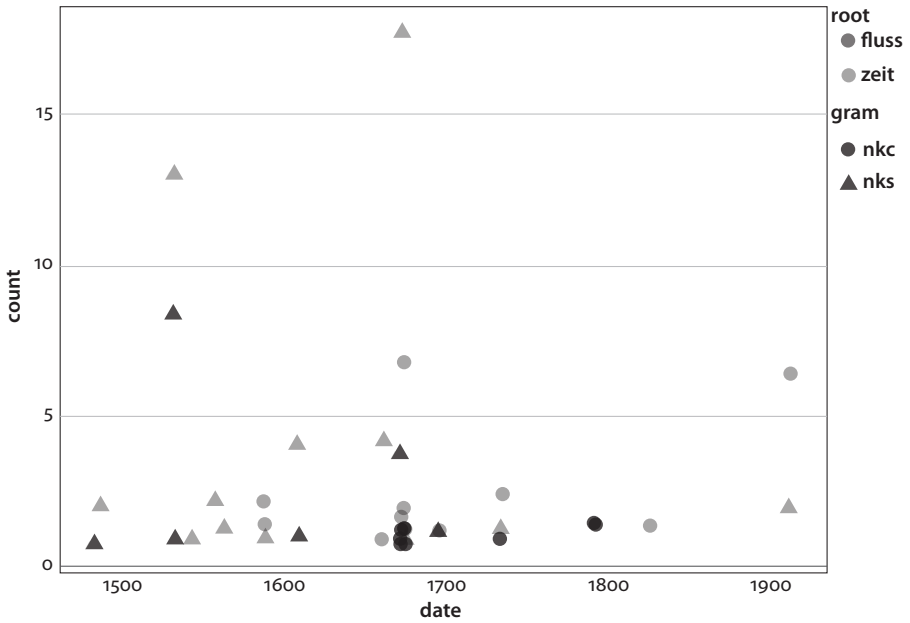
Figure 3. Distribution of the roots *zeit* ('time') and *fluss* ('flow') and their compounds (compound nouns *nkc*, simplex nouns *nks*)

More than half of references to menstruation in the corpus contain the root *zeit* or *fluss*. Both roots are used exclusively as simplex forms in the first half of the corpus period, and the frequency of compounds increases from the seventeenth century onwards. Derivations using these roots are not found. These findings can be interpreted in the context of contemporaneous meta-linguistic discussion about the appropriateness of German language for scientific use, which concerned its basic linguistic characteristics (Klein 2011a, 2011b). The possibility to form compound nouns productively was considered particularly valuable, as compounding was able to make up for the (perceived) lack of German words. Compounds could also denote complex concepts (cf. Schottelius 1663 as cited in Klein 2011b: 486).[20]

Is compounding then a strategy that is generally used when referring to menstruation? Table 7 shows the relative frequencies of simplex nouns compounds and derived nouns.

As shown in the table, references to menstruation tend to be morphologically simple. Derivations and compounds together account for less than 40% of the occurrences. Among the morphologically complex variants, compounds are more frequent than derivations. It is interesting to see that compounds occur in greater

---

20. For diachronic developments of German compounds, see Perlitz (2014) and Kopf (2018).

Table 7. Frequencies and proportions of morphological categories referring to menstruation: simplex nouns (*nks*), compound nouns (*nkc*) and derivation (*nkd*)

Gram	Count	Percent
nks	123	63%
nkc	46	23%
nkd	27	14%
M	196	100%

numbers only after having been emphasized by 17th-century grammarians as the most suitable linguistic means for the precision needed in scientific texts.

In addition to the morphological complexity of noun kernels, we map syntactic modifications of noun phrases: genitive attributes, adjective attributes, relative clauses and prepositional attributes. Less than 21% of the noun phrase references are syntactically modified or attributed in any grammatical sense. Figure 4 shows the modification and determination of nouns according to morphological categories (*nkd, nks, nkc*).
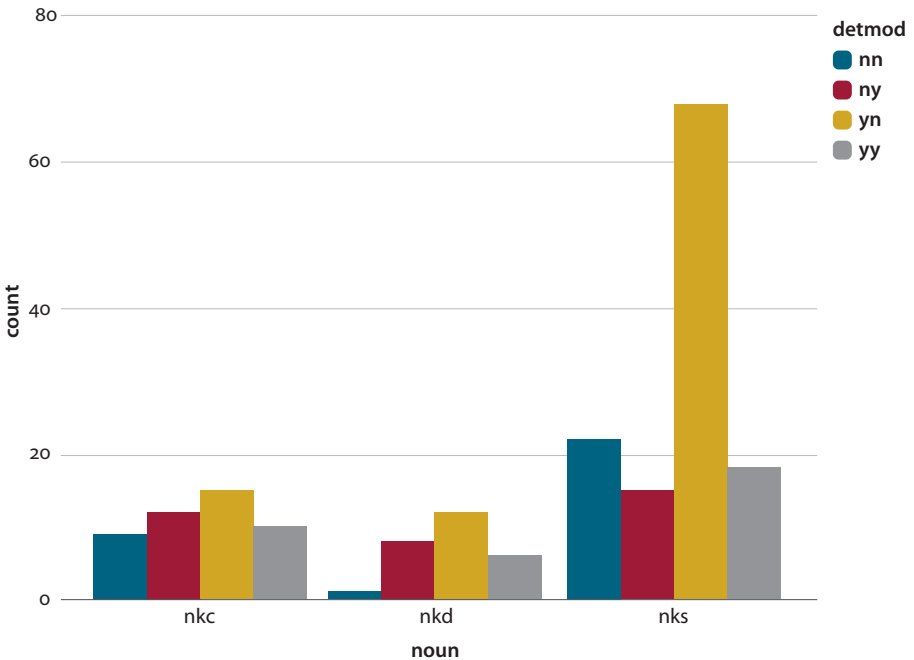
Figure 4. Frequency of determination and modification of the noun kernel. (*nn* = no modification and no determiner, *yn* = no modification but an overt determiner, *ny* = modification and no determiner, *yy* = modification and determiner)

As one would expect for Early New High German (ENHG) and New High German (NHG) (Ebert et al. 2013: 314–316), for all three types of noun kernels (nkc, nkd, nks) unmodified noun phrases with a determiner (detmod ="yn", e.g. *die Blumenzeit*) are the most frequent. Overall there seems to be no indication of a general tendency towards modification within the noun phrase. It is perhaps surprising that endo-centric compounds (which already have an internal modification) show the same number of external modification as simplex nouns: Examples like *der ůberflůſzigen Monat⸗Zeit* 'the superfluous monthly time' show that these modifications typically refer to pathological discrepancies to the normal state. Occurrences of bare nouns (detmod="nn") are highly marked (if not ungrammatical) in the German language of the period under discussion. The counts presented in Figure 3 are harder to interpret as we have no annotations of reference chains available and most of the occurrences are expected to be coordinated phrases with a dropped determiner (e.g. *zů irer*$_{DET}$ *blůmen oð []*$_{DET}$ ***feuchtikait*** 'to her flowers or []$_{DET}$ humidity').

To describe the structure of noun phrases in more detail, we look at the most frequent morphological type (*nks*). Figure 5 shows the distribution of modifications and determinations of simplex nouns over time.
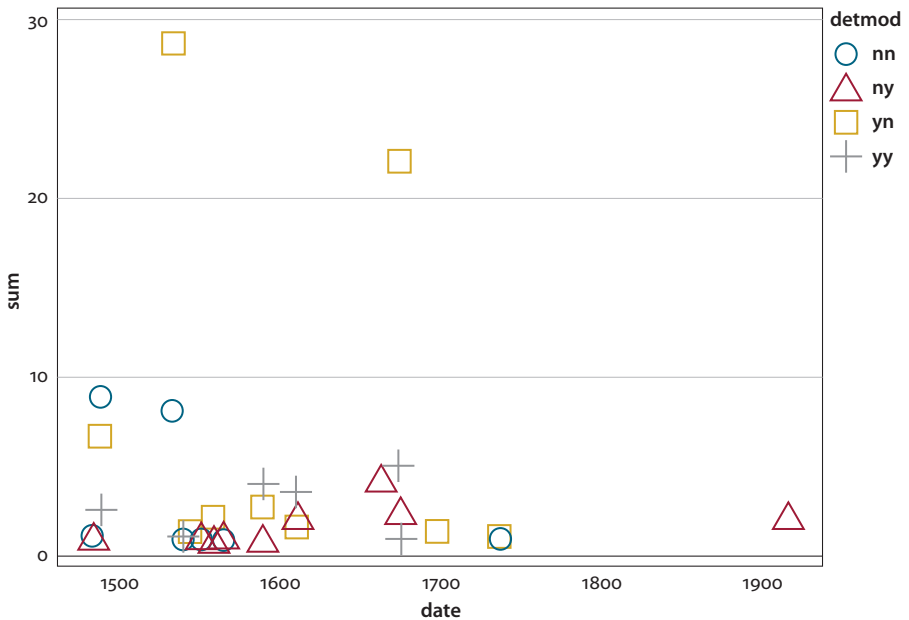
Figure 5. Frequency of modifications of simplex nouns (*nks*)

Simplex nouns tend to occur without determiners and modification in older documents (circles). Structures with a determiner and modifier, like *die verstopfte Blume* ('the congested flower'), tend to occur only before 1700 (plus signs). Overall, nouns with determiners and without modification constitute the majority of cases.[21] Even highly unspecific and ambiguous simple nouns such as *Zeit* do not necessarily need modification (or a determiner) to be used as references to menstruation.

4.3 Semiotic relation between denotation and terms

Our third set of research questions focuses on semantic relations between denotation and terms. How are the lexical, semiotic, morphological and syntactic variants distributed in our corpus across time? Which extra-linguistic factors could increase or decrease the use of certain strategies?

Figure 6 shows the diachronic distribution of morphological complexity and semiotic relations of noun kernels.
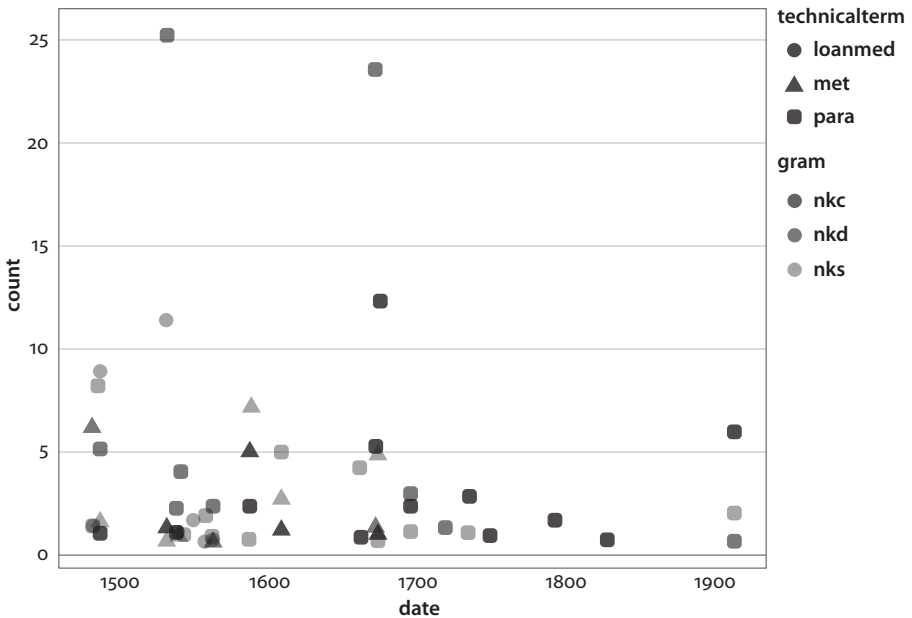
Figure 6. Frequency of noun kernels according to their grammatical annotation (colours, annotated in *gram*) and term classification (shapes, annotated in *technicalterm*)

---

21. There are two instances of nouns with a modifier but without a determiner from 1914. These nouns have a prenominal genitive such as *der Frauen Zeit* ('the women's time'). In some syntactic models the prenominal genitive would be analysed as a complex determiner. An alternative model would assume some form of syntactic blocking of the added determiner.

As can be seen, paraphrases (square) of any grammatical category are distributed over the entire corpus. The existence of grammatical variation in the references across 400 years shows no clear-cut change of strategies to create terminology. However, simplex nouns practically disappear from 1700 onwards.

Referring to menstruation using metaphorical expressions (as shown by the triangles in Figure 6) also ends rather abruptly in the second half of the seventeenth century. It is possible that this relates to the influential Enlightenment postulate of contentual interrelatedness in science. This postulate was introduced by Gottfried Leibniz in 1697 (Gardt 1998). Leibniz suggests that technical terms should be made up as far as possible from concrete words that cover observable properties of the referenced concept.

The *loanmed* reference type was used exclusively earlier in the period, with an accumulation of occurrences in the fifteenth and the first half of the sixteenth centuries. In addition, this type is attested in a single document in the eighteenth century (*Menses Mysterivm Sigillorvm* by Israel Hiebner, 1735). Despite the fact that Latin, the scientific *lingua franca*, was a supra-regional standard (cf. Wegera and Hartweg 1989; Lindner 2018 among many others), the *loanmed* category shows variation. Besides the technical term *menstruum*, also listed in reference works, there is also the lemma *menses*.[22] A qualitative analysis of the Latin technical term *menstruum* in this particular context shows that in the vast majority of cases, it does not appear as a sole reference to menstruation, but in combination with a vernacular expression. The use of the parenthesis *genannt menstruum* or *menstruum genannt* ('called menstruum') is typical (Example 3):

(3) *und also genutzt bringt auch die Krankheit der Frauen genannt menstruum*
     'and used in this way also brings the disease of women called menstruum'
                                            (Johannes von Cuba, *Gart der Gesundheit*, 1487)

References to medical phenomena through a juxtaposition of a vernacular and a Latin term is a common strategy around 1750 (Lindner 2018). This strategy can be seen as the linguistic manifestation of the attempt to combine two traditions of knowledge, the healers' tradition and the academic medical tradition, or as a strategy to highlight the author's expertise (Caflisch-Schnetzler 2018; Lindner 2018).

Some variants are found only in single text extracts. This is the case for derivations which are marginal. Their occurrences are distributed mainly over the first half of the period, the last example being from the early eighteenth century. Apart from this document, there are no derivations in the entire eighteenth and nineteenth centuries. The use of this morphological category is again attested only

---

22. *Menstruum* occurs 19 times (with 3 occurrences of the variant *menstrua*) and *menses* 5.

in the most recent corpus text from 1914, using the lexeme *Reinigung* ('cleansing'), as illustrated in Example (4).

(4) *ihr Geruch ist angenehmer und ihre Wirkung in Beschwerden der monatlichen Reinigung berühmter*
'its [the wormwoods] smell is more pleasant and its healing effect for the monthly cleansing is more famous'   (Friedrich Losch, *Kräuterbuch*, 1914)

How can this single occurrence be explained? Intertextuality could offer an explanation: compared to the specifically scientific texts on medicine and botany, the popular-scientific herbal book represents the development of the late medieval and early modern remedybook compendia combining herbology with medical knowledge (Taavitsainen 2018).[23] In the history of this genre, traditions of knowledge and typical linguistic patterns, including certain expressions, have played a role. Of interest here is the fact that the lexeme *Reinigung* ('cleansing') is documented once in a text from the sixteenth century (Adam von Bodenstein, *Wie sich meniglich*, 1557) and five times in a text from the seventeenth century (Bernhard Verzascha, *Theatrum Botanicum*, 1696). All instances of *Reinigung* from the seventeenth century text are modified with the adjective *monatlich* ('monthly') (Example 5):

(5) *Das destillierte Wermutwasser auf ein oder zwei Lot getrunken* […] *fördert die monatliche Reinigung der Weiber*
'The distilled wormwood water drunk on one or two lots […] promotes the monthly cleansing of the women'
(Bernhard Verzascha, *Theatrum Botanicum*, 1696)

Investigating the reception history of *Theatrum Botanicum* could reveal whether the author of *Kräuterbuch* had read it or relied on sources that belong to the same knowledge tradition.

5.   Summary and discussion

We have identified and annotated nominal references to menstruation in the RIDGES corpus. The variation and distribution of these references have been explored as an example of evolving medical terminology, which is an important aspect of the emergence of the German scientific register between 1550 and 1900. We have identified the following tendencies:

---

23. For a similar example in historical Hungarian medical texts, see Kuna (2017).

a. A decrease in formal variation: The early texts in the corpus have more linguistic variants than the later texts.
b. A decrease in semiotic variation: While the earlier texts contain many metaphorical and evaluative expressions, the later texts simply name the observable phenomenon.

Both tendencies provide evidence of progressive terminologisation. From a methodological point of view, we discussed the effects of corpus design. While it may seem obvious that terminologisation occurs hand in hand with the development of academic and scientific registers, only a corpus study with dense data points can provide a reliable basis for tracing such development. Our study thus goes beyond anecdotal evidence: all decisions have been documented and the findings can serve as a basis for further studies. To understand the processes of terminologisation in the history of a language, it is necessary to look at different terms; the corpus provides a good data source for this.

Extra-linguistic factors such as developments in the medical scientific community or metalinguistic discourse regarding the appropriateness of medical terms might have affected the choice for referring expressions to menstruation and are therefore also relevant in this context. This study illustrates that it is worthwhile to investigate both extra-linguistic social aspects – such as the development of knowledge and the formation of scientific communication networks – and the contemporary meta-linguistic discourse. In order to accurately assess the influence of these factors, they must be more precisely defined and modelled. Communication networks make an important contribution to the convergence of terms, as correspondence between researchers reduces variation. The role of standard works, such as the Zedler encyclopedia, and the institutionalisation of communication in academies and universities should be defined more precisely and their connection to language use needs to be modelled. The change from metaphorical and (imprecise, evaluative) phrasal references to more precise scientific terms can be seen in the context of the Enlightenment discourse supported by Leibniz on the requirements for a language of science, since paraphrased references correspond to the explicit postulate of *contentual interrelatedness*. The fact that the words *Fluss* 'flow' and *Zeit* 'time' can be found in most documents seems to contradict this. As discussed in Section 2.1, the reference identification of the expression cannot be determined without context and is therefore not biunique. This is illustrated in examples (6a–c).

(6) a. Walwurtz getruncken mit Wein / stellt den Frauen den Fluss Artzney Buchlein Der Kreutter, […]
         'Walnut root, if drunken with wine / initiates the women's flow'
                              (Johannes Tallat, *Artzney Buchlein der Kreutter*, 1532)

    b.   *und gestrichen um die Augen / nimmt den <u>Fluss</u> /*
        'd/ removes the flow'

<div align="right">(Johannes Tallat, Artzney Buchlein Der Kreutter, 1532)</div>

    c.   *dann hat er beide auf einem Berg in Italien (jedoch die ohne Geruch am Fuß des Bergs gegen Mitternacht, die wohlriechende hingegen in einem Wald desselben an einem <u>Fluss</u> gegen Mittag) und zu einer Zeit in ihrer Vollkommenheit gefunden.*
        'then he found both [herbs] on a mountain in Italy at the same time in their perfection (but to be more precise the one without smell he found at the foot of the mountain around midnight, the aromatic one, on the other hand, in a forest of the mountain, near a river, around noon)'

<div align="right">(Elisabeth Blackwell, Blackwellisches Kräuterbuch, 1750)</div>

In context, the noun *Fluss* ('flow') in (6a) is clearly identifiable as a reference to menstruation, whereas (6b) refers to another medical phenomenon (eye disease) and (6c) is even used in its primary meaning 'river'. Identifying the medical concept and the function of the term is not possible without the lexical context (*den Frauen*, 'to the women'; *um die Augen*, 'around the eyes'). The consistent occurrence of the nouns *Fluss* and *Zeit* nevertheless suggests conventionalisation, which can (but does not have to) be the consequence of the semiotic functionality of these lexemes. The applicability of the modern idea of terminology as a prescriptive standard for the precise communication within a scientific community must be adjusted for historical periods (and for the formation and evolution of terminology within new and established fields). The propagation of terminological items often starts from highly ambiguous everyday words or loan words and there may be a period of high variability (across registers and dialects). There is then a slow convergence period that culminates in one of the original words being selected as an unambiguous and neutral term.

Extra-linguistic factors would have to be incorporated into multi-dimensional language change models that take into account quantitative information. This can only be done using corpora or corpus tools. Due to limitations of the RIDGES corpus, we have only examined herbal texts, but it would be interesting to analyse other texts on gynaecological topics, e.g. medical texts or texts from the *Midwifery Corpus* (Whitt 2019 and this volume).

Our exploratory study has shown how essential the modelling and control of extra-linguistic context is for insights into diachronic developments. Although we could only identify tendencies, it is clear that modelling the contexts is possible. Since the development of terms is strongly linked to the differentiation of a particular social group or system, it would be an intriguing challenge to test hypotheses derived from the theory of language change and social systems (Zeige 2011) in the light of a validly modelled corpus-based study. Such a study could also have the potential to contribute to a better understanding of the mechanisms of language change.

References

Algee-Hewitt, Mark, and Mark McGurl. 2018. *Between Canon and Corpus: Six Perspectives on 20th-Century Novels*. Standford: Stanford Literary Lab. https://litlab.stanford.edu/Literary LabPamphlet8.pdf

Biber, Douglas. 1994. "Representativeness in Corpus Design." In *Current Issues in Computational Linguistics: In Honour of Don Walker*, ed. by Antonio Zampolli, Nicoletta Calzolari, and Martha Palmer, 243–257. Dordrecht: Springer. https://doi.org/10.1007/978-0-585-35958-8_20

Bode, Katherine. 2018. *A World of Fiction: Digital Collections and the Future of Literary History*. Ann Arbor: University of Michigan Press. https://doi.org/10.3998/mpub.8784777

Caflisch-Schnetzler, Ursula. 2018. "Die Bedeutung von Kommunikationsnetzwerken für die Entwicklung der deutschen Sprache im 18. Jahrhundert." In *Vernakuläre Wissenschafts-kommunikation. Beiträge zur Entstehung und Frühgeschichte der Modernen Deutschen Wissenschaftssprachen*, ed. by Michael Prinz, and Jürgen Schiewe, 87–100. Berlin, Boston: De Gruyter.

Ebert, Robert P., Oskar Reichmann, Hans-Joachim Solms, and Klaus-Peter Wegera (eds). 2013. *Frühneuhochdeutsche Grammatik*. Berlin: De Gruyter.

Evans, Jennifer. 2012. "'Gentle Purges corrected with hot Spices, whether they work or not, do vehemently provoke Venery': Menstrual Provocation and Procreation in Early Modern England." *Social History of Medicine* 25 (1): 2–19. https://doi.org/10.1093/shm/hkr021

Gardt, Andreas. 1998. "Sprachtheoretische Grundlagen und Tendenzen der Fachsprachenforschung." *Zeitschrift für Germanistische Linguistik* 26: 31–66.

Gläser, Rosemarie. 1995–1996. "Eigennamen in Wissenschafts- und Techniksprache." In *Namenforschung: Ein Internationales Handbuch zur Onomastik*, ed. by Ernst Eichler, 527–533. Berlin: De Gruyter.

Gloning, Thomas. 2007. "Deutsche Kräuterbücher des 12. bis 18. Jahrhunderts: Textorganisation, Wortgebrauch, funktionale Syntax." In *Gesund und krank im Mittelalter: Marburger Beiträge zur Kulturgeschichte der Medizin: 3. Tagung der Arbeitsgruppe "Marburger Mittelalter-Zentrum (MMZ)", Marburg, 25. und 26. März 2005*, ed. by Andreas Meyer, 9–88. Leipzig: Eudora-Verl.

Guentherodt, Ingrid. 1986. "Maria Cunitz und Maria Sibylla Merian: Pionierinnen der Deutschen Wissenschaftssprache im 17. Jahrhundert". *Zeitschrift für Germanistische Linguistik* 14 (1), 23–49. https://doi.org/10.1515/zfgl.1986.14.1.23

Hunston, Susan. 2009. "Collection Strategies and Design Decisions." In *Corpus Linguistics: An International Handbook*, ed. by Anke Lüdeling, and Merja Kytö, 154–168. Berlin, New York: Mouton de Gruyter.

King, Helen. 2015. "Menstruation." In *Oxford Classical Dictionary*. Oxford Research Encyclopedias. Oxford University Press. https://doi.org/10.1093/acrefore/9780199381135.013.4124

Klein, Wolf-Peter. 2011a. "Deutsch statt Latein! Zur Entwicklung der Wissenschaftssssprachen in der Frühen Neuzeit." In *Wissen Schaffen – Wissen Kommunizieren: Wissenschaftssprachen in Geschichte und Gegenwart*, ed. by Wieland Eins, Helmut Glück, and Sabine Pretscher, 35–47. Wiesbaden: Harrassowitz.

Klein, Wolf-Peter. 2011b. "Die Deutsche Sprache in der Gelehrsamkeit der Frühen Neuzeit: Vonder Lingua Barbarica zur HauptSprache." In *Diskurse der Gelehrtenkultur in der Frühen Neuzeit: Ein Handbuch*, ed. by Herbert Jaumann, 465–516. Berlin, New York: De Gruyter.

Kopf, Kristin. 2018. *Fugenelemente Diachron: Eine Korpusuntersuchung zu Entstehung und Ausbreitung der Verfugenden N+N-Komposita*. Berlin: De Gruyter. https://doi.org/10.1515/9783110517682

Koplenig, Alexander. 2019. "Against Statistical Significance Testing in Corpus Linguistics." *Corpus Linguistics and Linguistic Theory* 15 (2): 321–346. https://doi.org/10.1515/cllt-2016-0036

Krause, Thomas. 2019. ANNIS: A Graph-Based Query System for Deeply Annotated Text Corpora. PhD dissertation, Humboldt-Universität zu Berlin.

Krause, Thomas, and Amir Zeldes. 2016. "ANNIS 3: A New Architecture for Generic Corpus Query and Visualization." *Digital Scholarship in the Humanities* 31 (1): 118–139. https://doi.org/10.1093/llc/fqu057

Kruse, Britta-Juliane. 1999. *"Die Arznei ist Goldes wert": Mittelalterliche Frauenrezepte*. Berlin: de Gruyter. https://doi.org/10.1515/9783110885873

Kuna, Ágnes. 2017. "Die sprachlichen Muster von Anweisungen in ungarischen medizinischen Rezepten des 16.-17. Jahrhunderts." In *Sprachgeschichte und Medizingeschichte: Texte – Termini – Interpretationen*, ed. by Jörg Riecke, 195–208. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110524758-013

Lindner, Bettina. 2018. "Sprachenwechsel und Koexistenz: Zur Rolle des Deutschen und Lateinischen in Medizinischen Fallsammlungen des 17. und 18. Jahrhunderts." In *Vernakuläre Wissenschaftskommunikation. Beiträge zur Entstehung und Frühgeschichte der Modernen Deutschen Wissenschaftssprachen*, ed. by Michael Prinz, and Jürgen Schiewe, 255–270. Berlin, Boston: De Gruyter.

Lüdeling, Anke, Amir Zeldes, and Hagen Hirschmann. 2011. "Variationism and Underuse Statistics in the Analysis of the Development of Relative Clauses in German." In *Corpus Analysis and Diachronic Linguistics*, ed. by Yuji Kawaguchi, Makoto Minegishi, and Wolfgang Viereck, 37–57. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tufs.3.05lud

Odebrecht, Carolin, Malte Belz, Amir Zeldes, Anke Lüdeling, and Thomas Krause. 2017. "RIDGES Herbology: Designing a Diachronic Multi-Layer Corpus." *Language Resources and Evaluation* 51 (3): 695–725. https://doi.org/10.1007/s10579-016-9374-3

Odebrecht, Carolin, Laura Perlitz, Gohar Schnelle, and Catharina Fischer. 2020. *Annotationsrichtlinien zu Ridges Herbology Version 9.0*. Humboldt-Universität zu Berlin.

Perlitz, Laura. 2014. Konkurrenz zwischen Wortbildung und Syntax: Historische Entwicklung von Benennung. Bachelor's Thesis, Humboldt-Universität zu Berlin. https://doi.org/10.18452/14232

Pörksen, Uwe. 1998. "Deutsche Sprachgeschichte und die Entwicklung der Naturwissenschaften. Aspekte einer Geschichte der Naturwissenschaftssprache und ihrer Wechselbeziehung zur Gemeinsprache." In *Sprachgeschichte: Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*, ed. by Werner Besch, Anne Betten, Oskar Reichmann, and Stefan Sonderegger, 193–210. Berlin, New York: De Gruyter.

Reiner, Jörg L. 2003. *Terminologiesysteme als Grundlagen für die Semantische Interoperabilität von Heterogenen Anwendungssystemen sowie Deren Fachsprachlicher Zugang*. Berlin: Aka Akademische.

Reul, Christian, Dennis Christ, Alexander Hartelt, Nico Balbach, Maximilian Wehner, Uwe Springmann, Christoph Wick, Christine Grundig, Andreas Büttner, and Frank Puppe. 2019. "OCR4all – An Open-Source Tool Providing a (Semi-)Automatic OCR Workflow for Historical Printings." *Applied Sciences* 9 (22): 4853. https://doi.org/10.3390/app9224853

Riecke, Jörg. 2004. *Die Frühgeschichte der Mittelalterlichen Medizinischen Fachsprache im Deutschen: Bd 1: Untersuchungen. Bd 2: Wörterbuch*. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110895223

Roelcke, Thorsten. 2018. "Von der Wissenschaftlichen Eignung der Deutschen Sprache: Sprachreflexion in Barock und Aufklärung." In *Vernakuläre Wissenschaftskommunikation. Beiträge zur Entstehung und Frühgeschichte der Modernen Deutschen Wissenschaftssprachen*, ed. by Michael Prinz, and Jürgen Schiewe, 361–374. Berlin, Boston: De Gruyter.

Springmann, Uwe, and Anke Lüdeling. 2017. "OCR of Historical Printings with an Application to Building Diachronic Corpora: A Case Study Using the RIDGES Herbal Corpus." *Digital Humanities Quarterly* 11 (2). http://www.digitalhumanities.org/dhq/vol/11/2/000288/000288.html

Taavitsainen, Irma, and Päivi Pahta (eds). 2010. *Early Modern English Medical Texts: Corpus Description and Studies*. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.160

Taavitsainen, Irma. 2018. "Scholastic Genre Scripts in English Medical Writing 1375–1800." In *Diachronic Corpora, Genre, and Language Change*, ed. by Richard J. Whitt, 95–115. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/scl.85.05taa

Tiemann, Daniel; Schäffler, Arne and Willibald Pschyrembel. 2020. *Pschyrembel Klinisches Wörterbuch*. 268. revised edition, Boston: De Gruyter.

Wegera, Klaus-Peter, and Frédéric Hartweg. 1989. *Frühneuhochdeutsch: Eine Einführung in die deutsche Sprache des Spätmittelalters und der frühen Neuzeit*. Berlin: De Gruyter.

Whitt, Richard Jason. 2019. "Expressions of Knowledge in Early Modern English- and German-Language Midwifery and Gynaecological Texts (ca. 1500–1700): On the Use of Pronominal Subjects and *that*-Complement Clauses." *Neuphilologische Mitteilungen* 120 (2): 325–353.

Zedler, Johann Heinrich. 1731–1754. *Grosses vollständiges Universal Lexikon aller Wissenschafften und Künste*.

Zedler, Johann Heinrich. 1739. *Grosses vollständiges Universal Lexikon aller Wissenschafften und Künste: Ein und Zwantzigster Band Mi – Mt*. Leipzig.

Zeige, Lars Erik. 2011. *Sprachwandel und Soziale Systeme*. Hildesheim: Olms.

Changing functions, roles and representations

Language, labour and ideology

Constructing epistemologies of childbirth in the first three centuries of English-language midwifery texts (1540–1800)

Richard J. Whitt
University of Nottingham

Writings on midwifery and women's medicine related to childbirth reflect the many changes affecting this field during the Early Modern period, which in turn reflect changes in epistemological values Through the lens of critical discourse analysis, this paper discusses existing and emergent ideological demarcations between the wide range of midwifery texts published between 1500 and 1800 through a qualitative examination of prefatory material in these texts. It is shown that the earliest midwifery texts make much recourse to the authoritative knowledge of classical authors, following the tradition of medieval Scholasticism. As female midwives enter the textual scene, personal experience and empathy become valued sources of knowledge. Finally, the emergence of 'man-midwives' led to scientific prowess combined with personal experience touted as the most superior form of knowledge relating to the care of parturient women.

Keywords: Critical Discourse Analysis, Discourse-Historical Approach, midwifery, Scholasticism, ideology, textual practice

1. Introduction

The history of didactic midwifery texts over the past half-millennium is the history of a contested space, possibly more so than midwifery practice itself. Rights of access to a parturient woman and her body, as well as the ability to assist in the process of childbirth itself, rest on the social acceptance that certain groups of people possess the knowledge and capabilities necessary to participate in the welcoming of new life into the world. For most of human history, this has meant other women acting as midwives in assisting childbearing women in their travails. However, since the advent of birthing instruments and man-midwives in the seventeenth century gave rise to the increased medicalisation of childbirth, men have gained access to

this previously gynocentric space. In the realm of midwifery textbooks, however, medical men were advising midwives on best practice long before their presence in the birthing chamber became commonplace.

Tracing the ideological changes accompanying the early days of childbirth medicalisation – both in a practical and in a textual sense – is the focus of the current investigation, with an eye towards how the prefaces of some of the earliest vernacular English-language texts on midwifery and childbirth shed light on the epistemological changes that occurred from the 1540 publication of *The Byrth of Mankinde* (by Rösslin), the first English-language midwifery manual, through the explosion of texts about birthing practices written by practicing male surgeons ("man-midwives") during the latter half of the eighteenth century.

At the core of these ideological developments is that of a contested epistemological space: What types of knowledge (learned vs. experiential vs. scientific) relating to childbirth should be prioritised? Whose knowledge (midwife vs. surgeon) is most reliable or valuable in the birthing chamber? What constitutes an adequate evidence base for knowledge (experience and empathy vs. objective scientific reasoning)? Through the lens of Critical Discourse Analysis (CDA), this paper will trace how the shifting linguistic practices found in the prefatory material of midwifery manuals published between 1540 and 1800 both constitute and are constituted by the sociohistorical developments in both midwifery practice and changing attitudes about the medical care that ought to be afforded to expectant mothers.

Section 2 provides an overview of the history of midwifery and birthing practices from the early modern period onwards, as well as key developments in the types of didactic texts being published within this domain. Attention to the role of ideology in these shifting textual practices is in special focus. This is followed in Section 3 by a methodological overview with a focus on the texts under discussion and the Discourse-Historical and Critical Epistemic approaches to CDA receiving attention. The linguistic realisations of the most salient epistemological developments, as they are evidenced through the prefatory material of the midwifery manuals, are then discussed at length in Section 4. A section of concluding remarks and future prospects rounds out the discussion in Section 5.

2.   A brief history of midwifery and (textual) practice, 1540–1800

Until the birth of modern obstetrics, assisting with normal childbirth (i.e. a delivery proceeding with only minor or no complications) was an exclusively female undertaking throughout much of Europe. Only midwives and a few of the woman's closest female associates would be allowed to accompany the expectant mother into the birthing chamber. Fathers were never allowed into this space, and (male) surgeons

only came into the picture if there was an emergency, such as the extraction of a stillbirth. Although some ancient Western texts devoted to childbirth practice took an eye towards a female audience, most medieval writings on the subject – written in the tradition of Scholasticism – were aimed at a learned male audience as most women (including midwives) were illiterate during this period (Green 2008a: 29ff.). With the invention of the printing press in the fifteenth century and the subsequent vernacularisation boom across Europe, a number of scholars from a broad range of disciplines began to publish texts in local vernaculars such as English, German, French and Dutch rather than Latin, the heretofore language of learning and scholarship (see, for example, Green 2008a: 163ff.; Pahta and Taavitsainen 2010).[1] The first English-language text devoted to midwifery, *The Byrth of Mankinde*, was published by Richard Jonas in 1540. It was a translation of the Latin *De Partu Hominis* (1532), itself a translation of Eucharius Rösslin's German-language *Der Swangeren frawen vnd Hebammen roszgarten* ('The rose garden of pregnant women and midwives'), first published in 1513 and broadly considered to be the first midwifery manual published in a European vernacular (Arons 1994; Hobby 2009). Jonas' text was quickly superseded by Thomas Raynalde's 1545 translation, which corrected much of Jonas's erroneous translations of medical terminology and added an entirely new section based on Vesalian anatomy. Raynalde was a physician, whereas Jonas does not appear to have come from a medical background (see Hobby 2009). It was Raynalde's translation that enjoyed prominence for just over a century, going through a number of editions, until it was superseded in 1651 by the publication of Nicholas Culpeper's *A Directory for Midwives*.[2]

And contrary to any medieval texts on the subject, it is clear from the prefatory materials that both Rösslin's text and its English translations targeted a female audience, either pregnant women or midwives directly. Flügge's (1998: 364–367) discussion of Rösslin's text forming part of the statutory licensing examinations required of midwives in the German city-states implies that at least some midwives were literate by this period, while Richards' (2015) assertion that *The Byrth of Mankinde* was intended to be read out loud expands the targeted female audience of this text beyond literate midwives. Even so, neither Rösslin nor Jonas nor Raynalde ever found themselves in the birthing chamber, yet their works claimed an authoritative edge on best practice in childbirth. Part of this is due to the lingering

---

1. This vernacularisation of medical writing is what Pomata (2011a, b) describes as the advent of an "epistemic genre" – one focused on knowledge dissemination and exchange rather than the aesthetic or expressive purposes of medieval scholastic writings.

2. For more in-depth discussions of differences between Raynalde's and Jonas' translations, as well as differences between *The Byrth of Mankind* and its German source text, see Fissell (2004: 29–35), Hobby (2009) and Whitt (2018).

influence of medieval Scholasticism, one of whose tenets was that the best medical knowledge could be arrived at via a synthesis of the writings of classical (and some medieval) authors such as Hippocrates, Galen, Seranus, Avicenna and Rhazes. Indeed, much of Rösslin's text is a compilation of direct and indirect quotations from many of these authors' writings on midwifery and childbirth (Kruse 1994; Green 2009; Whitt 2018; cf. Taavitsainen 2001, 2012). This holds true for almost all other midwifery treatises published during the sixteenth and seventeenth centuries: they were either written by learned physicians who had never set foot in the birthing chamber, surgeons who only intervened in an emergency, or – as is the case with authors like Richard Jonas and John Sadler – learned men with little or no background in medicine at all.

It was not until 1609 that Louise Bourgeois, court midwife to Queen Marie de Médicis of France, penned *Observations diverses sur la stérilité, perte de fruit, et fécondité, accouchements, et maladies des femmes et enfants nouveaux nés* ('Diverse Observations on Sterility, Miscarriage, Fertility, Childbirth, and the Diseases of Women and Newborn Children'), that a midwifery treatise published by a practicing midwife involved in normal childbirth appeared. Consequently this is one of the first – if not the first – medical texts published in a European vernacular to be written by a woman (Perkins 1996; Lingo 2017; O'Hara 2017). Another court midwife, Justina Siegemund of the House of Brandenburg, was author of the first German-language midwifery treatise *Hoff-Wehe-Mutter* ('Court Midwife'), published in 1690 (although German translations of Bourgeois' text had been around since 1619).[3] Jane Sharp is generally believed to be the first woman/midwife to write an English-language midwifery text, *The Midwives Book* of 1671, although compared to Bourgeois and Siegemund, very little is known about Sharp biographically (Hobby 1999). This has led some scholars to go so far as to claim that Sharp was not a woman at all, but rather a male author writing with a female pseudonym to increase the marketability of his text (Walsh 2014; but cf. Hobby 1999). Evenden (2000) has gone so far to suggest that the first midwifery treatise to be published in English may actually precede Sharp's work by nearly two decades. In 1656, *The Compleat Midwifes Practice* was published by four authors referred to only by their initials (T. C., I. D., M. S. and T. B.), some of which correspond to midwives listed in the licensing records of the period (Evenden 2000: 8–9).[4] Aside from perhaps Sharp's text, all of these female-authored midwifery treatises share a belief that the extant midwifery texts written by men fall short both in terms of accuracy and a

---

3. For more information of Siegemund's text, see Gubalke (1985: 81ff.), Flügge (1998: 109ff.) and Tatlock (2005).

4. Catherine Turner (St Martin in the Fields, licensed 1632) and Dina Ireland (St Brides, licensed 1638).

necessary experiential frame of reference, and these women saw their works as necessary interventions to protect their fellow midwives from outdated or inaccurate advice.[5] But while women were entering the realm of medical writing, publication of midwifery texts written by men with substantially less experience in childbirth continued apace, most notably with the publication of Nicholas Culpeper's *A Directory for Midwives* in 1651 (itself also partly an amalgamation of earlier texts).[6]

Aside from the advent of female authors of midwifery treatises, the seventeenth century bore witness to another major development in the history of childbirth practices, one that no doubt set the gradual medicalisation of childbirth into motion: the use of instruments. This began with the Chamberlen family's secret use of forceps (among other instruments) around 1620. By the end of the century, instruments were catching on and their use became commonplace – although by no means universal – during the eighteenth century (Wilson 1995; Cody 2005: 31ff.; Lieske 2007–2009). Consequently, the "man-midwife", a male surgeon who would look after a pregnant woman from the early days of pregnancy through birth, began to displace the female midwife from the exclusively feminine space of normal childbirth. It should come as no surprise, then, that a number of these men published midwifery treatises of their own, but unlike the male authors of an earlier era, these men could draw on their own extensive experiences in childbirth practice rather than rely on experience restricted to medical emergencies and the accounts given either by their contemporaries or the authors of antiquity. Two of the most famous man-midwife authors from this period include Hendrik van Deventer (*The Art of Midwifery Improv'd*, a translation of his Latin text, was published in 1716) and William Smellie (his three-volume *A Treatise on the Theory and Practice of Midwifery* was published between 1752 and 1764). A number of women penned midwifery treatises during the eighteenth century, most notably Sarah Stone and her 1737 *A Complete Practice of Midwifery*, although they were small in number when compared with the number of men publishing during this period, especially in the latter half of the century.

---

5.   Indeed, Siegemund reports on two deaths resulting from midwives following Rösslin's advice against intervention when a placenta would not discharge as expected (Flügge 1998: 377).

6.   The works of Bourgeois and Siegemund were never translated into English during this period, even though the works of several non-English male surgeons were.

3.   Methodology

3.1   Knowledge, ideology and Critical Discourse Analysis

What should be clear from the brief account of changing textual practices in the field of midwifery during the early and late modern periods is that the value placed on certain types of knowledge plays a pivotal role in the sociocultural construal of who was deemed most qualified to advise others on midwifery practice, if not engage in this practice itself. The earliest authors of midwifery treatises placed great stock in the inherited wisdom of antiquity passed down through the medieval period to the present, whereas the female authors saw their own experiential knowledge as practicing midwives, in addition to empathy with their fellow women, as best placed in this field. Finally, the emergent man-midwife added scientific knowledge to the repertoire of experiential knowledge gained from increased involvement in normal childbirth. In short, these emergent textual and medical practices were the locus of one of the greatest epistemological conflicts in the history of medicine (Crawford 2015; Böhme 2017). Feminist critics (Cahill 2001; Sommers 2011; Staub 2011) have done well to bring the gendered nature of this conflict to the fore, although the most in-depth historical studies into the medicalisation of childbirth (Wilson 1995; Hanson 2004; McIntosh 2012) have highlighted that gender is just one among several variables at play here; economics, social status and women's self-agency have also been key factors driving developments in maternity care provision. As we shall see, broad brush strokes along the male-female divide fail to pick up on the nuanced ideological positions taken by the midwives and surgeons/man-midwives of this period. While some surgeons were keen on removing midwives from the purview of childbirth completely, others were intent on making midwives more informed about the nature of human anatomy and the technical processes at work in childbirth (after all, women at the time did not enjoy the same educational opportunities as men). Some midwives believed that beneficial relationships with surgeons could and should be established, while others took the position that childbirth should remain a woman-centred affair.[7]

While these developments in the history of midwifery are no secret, virtually no attention has been paid to the precise linguistic and textual practices that have accompanied the often discussed historical and medical changes in the field. Those that have devoted some attention to language usage (Keller 2000, 2003, 2007; Green

---

7.   Green (2008b) has made clear that, although normal childbirth remained a woman-centred phenomenon until the eighteenth century (probably because it was not even viewed as a medical issue), male medical practitioners were – throughout European history – regularly involved in treating women's health woes such as problems with menstruation and conditions of the breasts. Male-female medical encounters were thus not novel to the eighteenth century.

2008a: 251ff.) have done so without any systematic linguistic framework at hand, and the accompanying observations – while apt – lack in precision and technical rigour.[8] Critical Discourse Analysis (CDA) is chosen for the present analysis, for it combines a framework for linking salient ideologies present in discourse(s) to concrete linguistic realisations. Seeing as language is the junction of ideological practices – meaning(s) "in the service of power" (Fairclough 2010: 8; see also Thompson 1984) – as well as discursive practices – social actions manifest in the textual processes of production, distribution and consumption (Fairclough 2010: 56–68) – one can only presume that the language usage in these early midwifery texts will both constitute and be constituted by the shifting epistemologies surrounding the field. In particular, the Discourse-Historical Approach to CDA (Reisigl and Wodak 2001, 2016; Reisigl 2017) provides a robust framework allowing one to triangulate these ideological and discursive practices with their relevant sociohistorical context and subsequent operationalisation in language. In particular, Reisigl and Wodak (2001: 44–45, 2016: 33) highlight five salient discursive strategies indicative of an author's ideological positioning:

1. Nomination: How social actors, objects, phenomena, events, processes and actions are construed, mainly through the choice of nouns and verbs (e.g. *birth* vs. *hazardous passage, labouring woman* vs. *patient, man-midwife* vs. *boyish pretender*).
2. Predication: How social actors, objects, etc., are qualified –positively, negatively or neutrally – through the use of evaluative language (e.g. *delicate texture of the female constitution*, *tender* mother, *affectionate* wife, *barbarous* custom).
3. Argumentation: The justification (or questioning) of claims through the use of topoi or fallacies (e.g. *ad hominem* attacks against female midwives for their perceived ignorance of human anatomy).
4. Perspectivisation: Construal of the speaker's/writer's point of view and indications of involvement or distance (e.g. use of metaphors, mind styles [as expressed via, e.g., free indirect discourse], representations of [in]direct speech, etc.).
5. Intensification or Mitigation: Modifying the illocutionary force of an utterance (e.g. use of modality, hedges, metadiscourse, etc.).

There is obviously the possibility for overlap among these discursive strategies (the phrase *boyish pretender* involves both nomination and predication, for instance), but taken together, these strategies encompass a large range of linguistic and rhetorical phenomena that writers employ for ideological aims (more detailed descriptions of these strategies and accompanying language usage can be found in Reisigl and Wodak 2001, 2016, and Reisigl 2017).

---

8.   Pahta (2011) is one exception here, although her focus is restricted to theories of reproduction during the eighteenth century.

Since the focus here is on epistemological values and disputes found in early midwifery texts, van Dijk's (2011) framework of Critical Epistemic Discourse Analysis (CEDA) – a model of CDA that emphasises how knowledge itself can be the centre of ideological conflict – makes a perfect complement to Reisigl and Wodak's DHA (Discourse-Historical Approach). While it shares an interest in concrete linguistic realisations of discursive phenomena (2011: 37–39), CEDA focuses on the study of "the way knowledge is expressed, implied, suppressed, distributed, etc. in text and talk" and on how "general structures of power are related to such knowledge representation in discourse, for instance, what knowledge is emphasised or marginalised" (2011: 35–36; see also Potter 1996). This is precisely what the current study aims to show: how various and often conflicting attitudes towards different types of knowledge related to childbirth get expressed in the first three centuries of vernacular midwifery texts, and how these tie in with broader sociohistorical developments undergone by the discipline during this time period.

3.2   Creating a corpus of early midwifery texts (1540–1800)

The focus of the current discussion forms part of a larger project investigating language usage at the intersection of ideology, epistemology and midwifery practice from the sixteenth through the eighteenth century. Other issues of concern include changing metadiscursive textual practices through the three centuries in question (see, for example, Whitt 2018), the use of epistemic implicature surrounding key medical terminology (Plappert 2019), discursive use of modal and evidential markers (Taavitsainen 2001; Whitt 2016a, b), the rhetoric of the controversies surrounding the advent of man-midwives and instruments (Reinarz and Wynter 2015; Fritz et al. 2018), and point-of-view in narratives of childbirth (Simpson 1993). To this end, a corpus of complete midwifery texts, rather than mere extracts, representative of the first three centuries of vernacular medical writing and the accompanying changes in both textual and medical practice in the area of childbirth assistance has been compiled. For the sixteenth and seventeenth centuries, the *Early Modern English Medical Texts* corpus (Taavitsainen et al. 2010), or EMEMT, was used as a first port-of-call, and the complete version of all texts found in this corpus was accessed via *Early English Books Online – Text Creation Partnership* (EEBO-TCP).[9] Additional searches through both EEBO and the extant historical scholarship on midwifery during the early modern period were made to find additional texts for the corpus. For the eighteenth century, Lieske's (2007–2009) comprehensive 12-volume compendium on the history of eighteenth-century British

---

9.   EEBO URL: https://eebo.chadwyck.com/home; EEBO-TCP URL: https://quod.lib.umich.edu/e/eebogroup/.

midwifery proved invaluable, and all didactic midwifery treatises documented by Lieske were then accessed via *Eighteenth Century Collections Online* (ECCO).[10] Additional searches through ECCO and the historical scholarship were undertaken as well. Unfortunately, few of the relevant texts were available anywhere in machine-readable format, so PDF scans of the texts had to be run through Nuance OCR software to produce machine-readable files. This is less than ideal, seeing as OCR (optical character recognition) leaves behind both noise in the form of nonsense characters, as well as scanning errors. Fortunately for the texts involved, most of these errors are fairly straightforward and made consistently, so finding workarounds is possible.[11] Information about texts contained in the corpus as it currently stands can be found in the Appendix.[12]

Since the focus here is on ideologies concerning types of knowledge surrounding midwifery and childbirth practice in the prefatory materials of early and late modern midwifery manuals, relevant sections of each text were examined through close reading and manual analysis rather than via corpus searches using concordancing software. Hence the OCR issues discussed above were not so much an issue. This method was chosen because the linguistic realisation of the ideological nuances discussed above cannot easily be reduced to a fixed set of words or collocational patterns, and the prefatory materials were short enough (usually just a few hundred words, rarely over 1,000) that a detailed, "bottom-up" (Pahta and Taavitsainen 2010: 563) investigation was feasible. The discussion here is thus exclusively qualitative in nature.

---

10. ECCO URL: https://www.gale.com/intl/primary-sources/eighteenth-century-collections-online; ECCO-TCP URL: https://quod.lib.umich.edu/e/ecco/.

11. A good example of this is how the long-s, or *ſ*, is often rendered as *f* (less frequently as *t* and *l*) by OCR software. With this variation in mind, adjustments – such as the use of the wildcard * in corpus searches – can be made fairly straightforwardly.

12. Three names might seem conspicuously absent from this list: Percival Willughby, Chamberlen and Elizabeth Cellier. Willughby was a famous seventeenth-century man-midwife who penned *Observations in Midwifery* (King 1995; Evenden 2000: 50ff.; Keller 2003, 2007: 156ff.). However, this text cannot be found on either EEBO or ECCO because the text was not actually published en masse until the nineteenth-century edition, with editorial interventions by Henry Blenkinsop, was released. It was thus decided to exclude this text from the corpus for the time being. None of the Chamberlens actually published a treatise devoted to midwifery (although Hugh is the translator of Mauriceau's *The Diseases of Women with Child, and in Child-Bed* [1683], see Appendix; also note the absence of any of the Chamberlen publications in Wilson's 1995 study on *The Making of Man-Midwifery*). In a similar vein, Elizabeth Cellier – a midwife contemporary of Jane Sharp – also produced no midwifery treatise as such (King 1993; Cody 2005: 46ff.), and none of her publications were deemed suitable for the current investigation.

4.   An overview of language and ideology in the prefatory material

A key way of unpacking a text's ideological orientation is to examine how various "social actors" are represented (van Leeuwen 1995). This is often done through various strategies of nomination and predication (Reisigl and Wodak 2001, 2016; Reisigl 2017; see also van Leeuwen 1995). Coupled with this is how the key topics under discussion are depicted and discussed throughout the text(s) in question. Table 1 provides an overview of both the key actors involved in the world of childbirth mentioned in the prefatory material,[13] as well as the key themes discussed at length.

Table 1. The key actors and themes discussed in the prefatory material of sixteenth-, seventeenth- and eighteenth-century vernacular midwifery treatises

Key actors	Key themes
Women in labour	Knowledge (esp. bases of knowledge [experiential,
Female midwives	empathetic, textual] and whose knowledge is under
Man-midwives / Surgeons	discussion)
Physicians	Use of instruments
Newborns	Modesty and morality
Classical authors	
Contemporary authors	
Women (in general, as readers)	
Men (in general, as readers)	

Key here is unpacking the various and changing ideologies of the period by examining exactly how these actors are referred to (through the strategies of nomination and predication, for example) and the discursive construal of the key themes that recur throughout the first three centuries of vernacular midwifery writing.

4.1   The value and audience of knowledge dissemination

Perhaps the only thing all the authors of the midwifery treatises examined here have in common is that they all place a value on the dissemination of knowledge, and express a hope that their work will be beneficial to others. Of course some authors aimed at different types of audiences, while some valued different types of knowledge over other types, but the ultimate didactic aims of each text are underscored in every preface. Naturally the ideological space in which authors position themselves betray to some degree the sympathies of their respective audiences, at least insofar

---

13. Although the terms *preface* and *prefatory material* are used for consistency, sometimes an introductory chapter would serve the same function (laying out the author's motivations before entering into the technical details of physiology and birth). These types of chapters also factor into the current investigation.

as the target audience (market) of an author would most likely not engage with a text too far removed from its own ideological disposition. So in fact the ideologies present in these midwifery prefaces can gauge broader social values and practices related to childbirth as well.

Providing texts in local vernaculars is one impetus of some of the earliest authors of midwifery treatises. Thomas Phayer, in his preface to his translation of Jean Goeurot's *The Regiment of Life* (originally *L'Entretenement de vie*), explains his motives for making this text available in English:

(1)   […] but my purpose is here to doe them good that haue moost nede, yᵗ is to saye children: and to shewe the remedies that god hath created for the vse of mā, to distribut in Englishe to them that are vnlerned, part of yᵉ treasure that is in other lāguages, to prouoke them that are of better lernīg, to vtter theyr knowlege in such lyke attemptes […]

(Thomas Phayer, "The preface to the booke of children", in Jean Goeurot, *The Regiment of Life* [trans. by Thomas Phayer], 1550)¹⁴

Phayer's expressed desire is to do good for those that "haue moost nede", i.e. the children, by facilitating the education of the "vnlearned" through vernacular English text, as well improving the knowledge of the already learned. His focus here – through predication (Reisigl and Wodak 2001: 45ff.; cf. van Leeuwen 1995) – is on the attributes of specific key actors, namely newborns/children and birth assistants, although it is not immediately clear whether the unlearned are exclusively female midwives or include male surgeons as well. Given that in 1550, the only other vernacular midwifery text available would have been *The Byrth of Mankinde*, Phayer certainly had reason to believe his translation of Goeurot's work would find a receptive audience due to its novelty in any case. John Sadler targeted his 1636 *The Sick Womans Priuate Looking-glasse* more overtly towards women, whom he believed to be in desperate need of advice concerning their reproductive systems:

(2)   BECAVSE I had my being from a woman, I thought none had more right to the grape than she which planted the vine. Considering therefore the manifold distempers of body, which yee Women are subject unto through your ignorance & modestie, I could not but doe my best, to informe and advise you in the conservation of your own health. And when I had spent some meditations, and consulted with *Galen* and *Hippocrates* for my proceeding; amongst all diseases incident to the body, I found none more frequent, none more perilous then those which arise from the ill affected wombe […]

(John Sadler, *The Sick Womans Priuate Looking-glasse*, 1636)

---

14. Some texts did not paginate their prefatory materials, so precise page numbers are not always available for the examples given.

Expressing a commonly held belief of the time, Sadler ascribes nearly all of women's malaise to their "ill affected" wombs which, when combined with their own "ignorance & modestie", place women's health ever on the edge of peril. Throughout Extract (2), women and their bodies are frequently passivised, either grammatically ("are subject unto", "ill affected") or through the nominalisation of actions ("conservation of your own health"). Best-placed to advise these women are authors such as Sadler (who himself did not have a background in medicine), whose extensive knowledge of the authors of antiquity make him an authority on gynaecological matters; therefore, he is capable of taking a (grammatically and socially) active role in advising the passive sex (see van Leeuwen's [1995: 42ff.] discussion of discursive role allocation, which includes phenomena such as passivisation and activation). A century later, when Sarah Stone published *A Complete Practice of Midwifery* (1737), the express goal of knowledge dissemination to and for women was also foremost in Stone's mind, although her tone is quite different:

(3)   THE Occasion of my publishing this small Treatise is, in hopes it may prove instructive to some Women Professors in the Art of Midwifery; and inform them in a right, safe, and just practice of that Art: that they may be able to deliver in difficult Labours, as well as those that are not so. For I cannot comprehend, why Women are not capable of compleating this business when begun, without calling in of Men to their assistance, who are often sent for, when the Work is near finish'd; and then the Midwife, who has taken all the pains, is accounted of little value, and the young men command all the praise. Which unskilful practices of Women-Midwives being often repeated, give occasion for Pregnant Women to bespeak them, so that is is become quite a fashion; especially with the *Bristol* Ladies.

(Sarah Stone, *A Complete Practice of Midwifery*, 1737: ix–x)

Stone's audience is clearly narrower than Phayer's but similar to Sadler's. But unlike Sadler, women take a noticeably active role in the grammar, and in the discourse ("women professors", "they may be able to deliver", "women […] compleating this business", "taken all the pains", "to bespeak them", etc.). And although both Phayer and Stone expressly wish their work to better inform childbirth practices to make passage into the new world as safe as possible for the neonate (whereas Sadler's focus was more on women's health in general, only part of which encompassed childbirth), Stone expressly focuses her attention on the need for female midwives to better inform themselves. This is to abate the encroachment of male practitioners – who Stone labels as *boyish/Gentlemen/young* and *well-assur'd pretenders* throughout her preface (another use of nomination and predication strategies) – into the sphere of normal childbirth. Such an intrusion often results in the men-midwives taking credit for much of the work already done by the female

midwives, thus "command[ing] all the praise" for the former and their profession while diminishing the capacity and credibility of the latter, who are "accounted of little value". Women are thus left in a passive state yet again. The stance of Stone's "pretenders" can be illustrated by George Counsell, author of *The Art of Midwifery* (1752), who takes quite a different view on the matter:

> (4)   AS this following Treatise was chiefly designed for the Use of Midwives, and such Practitioners in England, as are unacquainted with any other Language than their own; I have taken care to write in a plain, easy Stile, such as I apprehend will render it of more Use and Benefit to them than most Books I have met with upon the same Subject: But my chief Care and Concern has been, throughout the Whole, to lay down in the plainest, and at the same time the most concise Terms, a Method of Practice, which, for many Years past, I can safely aver, upon my Conscience, has proved successful, and in many Cases far beyond my own Expectation: And should I meet with the Approbation of the Candid and Learned in the Faculty of Physic, and of other Gentlemen of Experience in this Profession, I shall think myself happy; but much more so, should I ever hear, that I have saved the Life of one single Woman or Infant, by the Doctrine herein laid down, or the cautious Admonitions I have given to young Practitioners. And, certainly, such Cautions and Admonitions were never more wanted than at this present Juncture: For it is a Truth too well known, that Mothers and their Children are daily, if not hourly, destroyed [such is the Practice of Midwifery in our Days] by ignorant Wretches, in almost every State of Life, a Pack of young Boys, and old superannuated Washer-women, who are so impudent and so inhuman as to take upon them to practise, even in the most difficult Cases, which as possibly occur.
>
> (George Counsell, *The Art of Midwifery*, 1752:ix–xi)

On the one hand, Counsell shares Phayer's view of vernacular medical writing and the possibility of reaching as wide an audience as possible, but he differs from Stone as to who is responsible for problems in the current practice of childbirth assistance, as well as who is best positioned to improve the situation (and be the targeted recipient of his knowledge transfer). On the other hand, he shares with Stone the frequent employment of nomination strategies to both elevate and denigrate. But whereas Stone blames "Gentlemen pretenders" (man-midwives) for many of the problems she has dealt with, Counsell seeks the approval of these same "Gentlemen of Experience" and rather blames "ignorant Wretches", a "Pack of young Boys" (inexperienced man-midwives) and "superannuated Washer-women" for problems of botched births and infant mortality. These nominations also reveal what van Leeuwen (1995: 54) refers to as identification, the marking of social actors "in terms of […] what they are".

This is in contrast to functionalisation, whereby social actors are defined in terms of what they *do*. All of Counsell's labels thus obscure the issue of midwifery practice (whether competent or not) and rather take a more personal aim at "impudent and inhuman" individuals. He then advocates for stricter licensing of (female) midwives and restricts the use of instruments to his fellow man-midwives. That said, his text is still at least partly targeted at female midwives in the hopes of improving their knowledge as well (the Wretches and Washer-women are generally believed to express his distaste for women of a lower economic standing acting as impromptu midwives rather than professional female midwives, see Vol. 9 [2009] of Lieske 2007–2009: 175–177). Similar but arguably more charitable goals were expressed by Brudenell Exton the year before:

> (5)   If Midwives will attentively read this Treatise, I hope it will be of great Service to them in regulating their Practice, as well as making them sensible what is their proper Business, and what not. I have always observed, that the more Knowledge they have, the readier they are to send for timely Assistance, in Cases of Danger: For it must be the greatest Ignorance that occasions them to keep Women under their Hands many Hours, by giving them fallacious Hopes, when they perhaps have it not in their Power to give them the least Assistance.
> (Brudenell Exton, *A New and General System of Midwifery*, 1751: 11–12)

Exton here explicitly targets his text at female midwives, whom he implicitly suggests are perfectly suitable to assist with normal childbirth. It is when complications arise that the (male) surgeon is to be summoned (historically, this has always been the case). Exton desires his treatise to inform women exactly *when* this was to be done. Women do take an active role in the discourse, yet through a series of nominations and predications ("proper Business", "timely Assistance", "greatest Ignorance", "fallacious Hopes") he implies that these midwives are insensible and unaware of their "proper" place in the medical hierarchy. They nonetheless form a part of his target audience. Some of the latter nineteenth-century male authors exclude women altogether from their focus. In the prefaces to Alexander Hamilton's *Outlines of the Theory and Practice of Midwifery* (1787) and Thomas Denman's *An Introduction to the Practice of Midwifery* (1794), for example, midwifery is still considered a practice in its own right, yet the target audience is exclusively male surgeons. The discursive erasure of female midwives suggests that women have completely fallen out of the scope of these men's conception of legitimate birth attendants (see van Leeuwen's [1995: 38ff.] discussion of exclusion and the suppression/backgrounding of social actors; cf. van Dijk 1995).

4.2   Authors as the curators of knowledge

We now turn to an examination of how various authors positioned themselves epis-temologically, arguing their own knowledge and experiences were superior to other competing forms at the time (see discussions on "perspectivisation" in Reisigl and Wodak 2001, 2016, and Reisigl 2017). In Example (2) we saw how Sadler makes recourse to the works of Galen and Hippocrates in an effort to bolster his credi-bility, as the knowledge of the ancient authors was highly valued in learned circles. If we take Evenden's (2000) lead that *The Compleat Midwife's Practice* was the first English-language midwifery treatise published by women, we can see that practicing midwives place value in a different form of knowledge:

(6)   Now Christian Reader, to give thee a true information of what we have here done for thy good, we shal not only Justifie from our own experiences, but fully demonstrate from the writings of the best practises, both of the *French, Spanish,* and *Italians,* and other Nations; and we must cleerly confesse, that we are highly obliged to the incomparable labours of that most famous woman of the world, *Madam Long Bourgeo,* late Midwife to the Queen of *France,* the praises that we read of all those that ever heard of her, are not so much a flourish as truth, for her reasons are solid experiences, and her witnesses have been all of the most emminent persons of *France* […]

(T. C. et al., *The Compleat Midwife's Practice*, 1656)

For one, these women place much stock in their own experiences in midwifery prac-tice, something none of their male contemporaries could claim, as the man-midwife was not yet commonplace. Secondly, Louise Bourgeois and her writings are valued as key sources of information rather than the work of ancients or a male contempo-rary. It is worth noting that no male-authored treatise of this period makes explicit recourse to Bourgois' work, despite her extensive experience in the field. In fact, T. C. et al. draw attention to the deficiency of extant midwifery texts, focusing spe-cial attention on Nicholas Culpeper's *Directory for Midwives* (Part 1, 1651; Part 2, 1662), condemned as "desperately deficient" and based on highly flawed material (i.e. other male-authored midwifery and medical manuals). Knowledge gained from personal experience, and based on the first-hand experience of others, super-sedes more mediated learned knowledge here. Taking this a step further, Elizabeth Nihell believes women qua women are the sole proprietors of such knowledge and should have exclusive rights to the birthing chamber. Men-midwives are to play no part here, and those men who inject themselves into this process do so out of dubious motives and with imperfect knowledge:

(7)   I might plead that of so many authors who have, with the utmost confidence and the utmost absurdity, written upon the art of midwifery, without understanding any thing at all of it. The truth is, that my very natural and strong attachment to the profession, which I have long exercised and actually do exercise, created in me an unsuppressible indignation at the errors and pernicious innovations introduced into it, and every day gaining ground, under the protection of Fashion, sillily fostering a preference of men to women in the practice of midwifery: a preference first admitted by credulous Fear, and admitted without examination, upon the so suspicious recommendation of those interested to make that Fear subservient to their selfish ends.

(Elizabeth Nihell, *A Treatise on the Art of Midwifery*, 1760:ii)

Indeed, Nihell's entire treatise is not so much focused on midwifery practice as such, but rather on an extended exposition against the phenomenon of man-midwifery (see Vol. 6 [2008] of Lieske 2007–2009: viiff.). Unlike earlier generations of male authors, though, the targets of Nihell's invective are not men without practical knowledge, but medical men who lack the lived and embodied experience that only women can have. The men-midwives of her time saw – and even conceptualised – their own experience in a fashion similar to that of T. C. et al. in Example (6) above:

(8)   I shall conclude this tedious, and (as it may appear to some) impertinent Preface, with observing, that the Contents of the subsequent Chapters are the result of Experience in the Disorders therein mentioned, and that I submit, whatever is advanced in either of them, or the preceding ones, with all due Reverence, to the Consideration of the Publick, humbly hoping, that my Want of Years, and consequently of long Experience, will in some measure atone for the many Faults which occur in this Performance, and that Maturity of Age will enable me to present it with a Work more worthy its accepting.

(Giles Watts, *Reflections on Slow and Painful Labours*, 1755:x)

Both T. C. et al. and Watts make recourse to their "long Experience" as the source of their knowledge and position themselves as best placed to provide advice on childbirth assistance, accentuated by frequent use of first-person pronouns throughout their texts. Watts even adds "Maturity of Age" to his list of qualifications, embedded in a plethora of self-deprecations – admittedly quite common in the prefaces of the period – and deference to the likes of William Smellie (vis-à-vis T. C. et al.'s esteem for Louise Bourgeois).[15] By the end of the eighteenth century, some man-midwives

---

15.   This comes earlier in Watts' Introduction: "And with respect to the Obstetretick Art, which is now in an eminent Manner improved, and that chiefly by the indefatigable Application of the great Dr. SMELLIE, general Treatises are become, in a Manner, altogether unnecessary" (1755: vi). Even so, Watts believes his "long Experience" and "Maturity of Age" enable him to make some additional contributions, however modest, to the extant body of knowledge available in the ever increasing number of midwifery treatises being published during this period.

had come to view childbirth assistance as completely falling within the purview of professional medical practice, thus excluding women from the picture completely. Through a crafty (re)definition of the term *midwifery* itself, Alexander Hamilton, who made no mention of women in his preface, positions the knowledge and experience required for midwifery squarely within that of the scientifically-informed, androcentric world of surgery and physic. Female midwives have become completely "suppressed" (van Leeuwen 1995: 39) social actors who seem to no longer play any significant role in childbirth assistance:

> (9) Midwifery, which may be defined "The art of facilitating the birth of children," is to be considered in much the same light as the other parts of surgery. Theory is less essential to it, as it chiefly consists in an operation which requires a dexterity, only to be learned by practice. But, taken in a more enlarged sense. Midwifery may be defined, "The art of facilitating the birth of children, and of managing pregnant and puerperal women." A part of it, therefore, has still a relation to the practice of physic: and, as such, must be involved in the same difficulties and obscurities.   (Alexander Hamilton, *Outlines of the Theory and Practice of Midwifery*, 1787:xiv)

Thus began the medicalisation of normal childbirth. Not only are male surgeons/ man-midwives now best placed to provide assistance before, during and after childbirth, they have also become the "epistemic elites" (van Dijk 2011: 40; cf. van Dijk 1995), dominating the means of knowledge production and dissemination through their publications.

5.   Concluding remarks

The discussion here has provided a broad overview of the ideological developments in the history of midwifery, as evidenced in the English-language midwifery treatises printed between 1540 and 1800. While female midwives had long enjoyed near-exclusive rights to assist in the birthing chamber, male authors positioned themselves as epistemologically superior to these women as long as these treatises were in print, even though practical, hands-on experience often never entered into the equation. It was only with the advent of the man-midwife that experience alongside scientific knowledge became central tenets of expertise in the discipline, even though female midwives since Louise Bourgeois attempted to use their expertise to garner credibility for their respective texts.

Throughout the first three centuries of midwifery writing, both male and female authors attempted to discursively construe a textual space that privileged the types of knowledge most conducive to their own version of midwifery practice. Absent from the current discussion is an in-depth examination of other key issues at play

during this period: namely, the role of modesty and morality in the construal of a legitimate and qualified midwife, as well as the use of instruments in midwifery practice. The earliest midwifery treatises (by both men and women) put much stock in moral character as a necessary precursor to being a capable midwife, whereas the man-midwives rarely if ever considered their own moral dispositions as relevant to their practice: all that supposedly mattered was technical/scientific knowledge and experience. The use of instruments proceeds along less gendered lines, as both pro- and anti-movements consisted of both men and women. In addition, religious and political affiliation were just as likely to be deciding factors in one's position on the matter as gender was (see, for example, Harley 1993, Wilson 1995, Lieske 2007–2009, King 2012).

What I hope to have demonstrated is the nuanced ideological developments in this field, and how these can be accessed linguistically. Although my approach has been qualitative in nature, more quantitative corpus techniques can be employed on entire texts to get a more concrete overview on issues like the frequency with which instruments such as the forceps and crochet were discussed with a positive, neutral, or negative shading, and by whom, as well as the degree to which certain types of epistemic qualification (e.g. through the use of modality or evidentiality) or implicature were employed across temporal, gendered and ideological parameters.

Now that the macro-level ideological discursive landscape surrounding these midwifery texts has been surveyed, we are in a better position to understand any relevant microscopic linguistic details emerging from subsequent investigations. There is still much to be learned from the intricate and nuanced history of midwifery, especially concerning the changing discursive practices found in the writings devoted to a discipline which remained at the fringes of institutionalised medicine and medical history until only fairly recently.

References

Arons, Wendy. 1994. "Translator's Introduction." In *When Midwifery Became the Male Physician's Province: The Sixteenth Century Handbook* The Rose Garden for Pregnant Women and Midwives, *Newly Englished*, by Eucharius Rosslin, trans. by Wendy Arons, 1–25. Jefferson, NC: McFarland & Company.

Böhme, Gernot. 2017. "Midwifery as Science: An Essay on the Relation between Scientific and Everyday Knowledge." In *Society and Knowledge: Contemporary Perspectives in the Sociology of Knowledge and Science*, ed. by Nico Stehr, and Volker Meja, 373–392. New Brunswick: Transaction Publishers. https://doi.org/10.4324/9781315129884-17

Cahill, Heather A. 2001. "Male Appropriation and Medicalization of Childbirth: An Historical Analysis." *Journal of Advanced Nursing* 33 (3): 334–342. https://doi.org/10.1046/j.1365-2648.2001.01669.x

Cody, Lisa Forman. 2005. *Birthing the Nation: Sex, Science, and the Conception of Eighteenth-Century Britons*. Oxford: Oxford University Press.

Crawford, Patricia. 2015. *Blood, Bodies and Families in Early Modern England*. London: Routledge. https://doi.org/10.4324/9781315838434

van Dijk, Teun A. 1995. "Discourse, Power and Access." In *Texts and Practices: Readings in Critical Discourse Analysis*, ed. by Carmen Rosa Caldas-Coulthard, and Malcolm Coulthard, 84–104. London: Routledge.

van Dijk, Teun A. 2011. "Discourse, Knowledge, Power and Politics: Towards Critical Epistemic Discourse Analysis." In *Critical Discourse Studies in Context and Cognition*, ed. by Christopher Hart, 28–63. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/dapsac.43.03van

Evenden, Doreen. 2000. *The Midwives of Seventeenth-Century London*. Cambridge: Cambridge University Press.

Fairclough, Norman. 2010. *Critical Discourse Analysis: The Critical Study of Language*. London: Routledge.

Fissell, Mary E. 2004. *Vernacular Bodies: The Politics of Reproduction in Early Modern England*. Oxford: Oxford University Press.

Flügge, Sibylla. 1998. *Hebammen und heilkundige Frauen: Recht und Rechtswirklichkeit im 15. und 16. Jahrhundert*. Frankfurt a. M./Basel: Stroemfeld.

Fritz, Gerd, Thomas Gloning, and Juliane Glüer (eds). 2018. *Historical Pragmatics of Controversies: Case Studies from 1600 to 1800*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/cvs.14

Green, Monica H. 2008a. *Making Women's Medicine Masculine: The Rise of Male Authority in Pre-Modern Gynaecology*. Oxford: Oxford University Press.

Green, Monica H. 2008b. "Gendering the History of Women's Healthcare." *Gender & History* 20 (3): 487–518. https://doi.org/10.1111/j.1468-0424.2008.00534.x

Green, Monica H. 2009. "The Sources of Eucharius Rösslin's 'Rosegarden for Pregnant Women and Midwives' (1513)." *Medical History* 53: 167–192. https://doi.org/10.1017/S0025727300000193

Gubalke, Wolfgang. 1985. *Die Hebamme im Wandel der Zeiten: Ein Beitrag zur Geschichte des Hebammenwesens*. 2nd edition, ed. by Ruth Kölle. Hannover: Elwin Staude.

Hanson, Clare. 2004. *A Cultural History of Pregnancy: Pregnancy, Medicine and Culture, 1750–2000*. Basingstoke: Palgrave. https://doi.org/10.1057/9780230510548

Harley, David. 1993. "Provincial Midwives in England: Lancashire and Cheshire, 1660–1760." In *The Art of Midwifery: Early Modern Midwives in Europe*, ed. by Hilary Marland, 27–48. London: Routledge.

Hobby, Elaine. 1999. "Introduction." In *The Midwives Book, or the Whole Art of Midwifery Discovered*, by Jane Sharp, ed. by Elaine Hobby, xi–xxxi. Oxford: Oxford University Press.

Hobby, Elaine. 2009. "Introduction." In *The Birth of Mandkind: Otherwise Named, The Woman's Book*, by Thomas Raynalde, ed. by Elaine Hobby, xv–xxxix. Farnham: Ashgate.

Keller, Eve. 2000. "Embryonic Individuals: The Rhetoric of Seventeenth-Century Embryology and the Construction of Early-Modern Identity." *Eighteenth-Century Studies* 33 (3): 321–342. https://doi.org/10.1353/ecs.2000.0027

Keller, Eve. 2003. "The Subject of Touch: Medical Authority in Early Modern Midwifery." In *Sensible Flesh: On Touch in Early Modern Culture*, ed. by Elizabeth D. Harvey, 62–80. Philadelphia: University of Pennsylvania Press.

Keller, Eve. 2007. *Generating Bodies and Gendered Selves: The Rhetoric of Reproduction in Early Modern England*. Seattle: University of Washington Press.

King, Helen. 1993. "The Politick Midwife: Models of Midwifery in the Work of Elizabeth Cellier." In *The Art of Midwifery: Early Modern Midwives in Europe*, ed. by Hilary Marland, 115–130. London: Routledge.

King, Helen. 1995. "'As if none understood the art that cannot understand Greek': The Education of Midwives in Seventeenth-Century England." In *The History of Medical Education in Britain*, ed. by Vivian Nutton, and Roy Porter, 184–198. Amsterdam/Atlanta: Rodopi.

King, Helen. 2012. "Midwifery, 1700–1800: The Man-Midwife as Competitor." In *Nursing and Midwifery in Britain since 1700*, ed. by Anne Borsay, and Billie Hunter, 107–127. Basingstoke: Palgrave Macmillan. https://doi.org/10.1007/978-1-137-02421-3_5

Kruse, Britta-Juliane. 1994. "Neufund einer handschriftlichen Vorstufe von Eucharius Rößlins Hebammenlehrbuch *Der schwangeren Frauen und Hebammen Rosengarten* und des *Frauenbüchleins* Ps.-Ortolfs". *Sudhoffs Archiv* 78: 220–236.

van Leeuwen, Theo. 1995. "The Representation of Social Actors." In *Texts and Practices: Readings in Critical Discourse Analysis*, ed. by Carmen Rosa Caldas-Coulthard, and Malcolm Coulthard, 32–70. London: Routledge.

Lieske, Pam. 2007–2009. *Eighteenth-Century British Midwifery*. 12 Vols. London: Pickering & Chatto.

Lingo, Alison Klairmont. 2017. "Editor's Introduction." *Midwife to the Queen of France:* Diverse Observations, by Louise Bourgeois, trans. by Stephanie O'Hara, and ed. by Alison Klairmont Lingo, 1–65. Toronto: Iter Press.

McIntosh, Tania. 2012. *A Social History of Maternity and Childbirth: Key Themes in Maternity Care*. London: Routledge.

O'Hara, Stephanie. 2017. "Translator's Introduction." *Midwife to the Queen of France:* Diverse Observations, by Louise Bourgeois, trans. by Stephanie O'Hara, and ed. by Alison Klairmont Lingo, 67–77. Toronto: Iter Press.

Pahta, Päivi. 2011. "Eighteenth-Century English Medical Texts and Discourses on Reproduction." In *Languages of Science in the Eighteenth Century*, ed. by Britt-Louise Gunnarsson, 333–355. Berlin/Boston: de Gruyter.

Pahta, Päivi, and Irma Taavitsainen. 2010. "Scientific Discourse." In *The Handbook of Historical Pragmatics*, ed. by Andreas H. Jucker, and Irma Taavitsainen, 549–586. Berlin/New York: de Gruyter. https://doi.org/10.1515/9783110214284.7.549

Perkins, Wendy. 1996. *Midwifery and Medicine in Early Modern France: Louise Bourgeois*. Exeter: University of Exeter Press.

Plappert, Garry. 2019. "Not Hedging but Implying: Identifying Epistemic Implicature through a Corpus-Driven Approach to Scientific Discourse." *Journal of Pragmatics* 139: 163–174. https://doi.org/10.1016/j.pragma.2018.09.001

Pomata, Gianna. 2011a. "Observation Rising: Birth of an Epistemic Genre, 1500–1650." In *Histories of Scientific Observation*, ed. by Lorraine Daston, and Elizabeth Lunbeck, 45–80. Chicago: University of Chicago Press.

Pomata, Gianna. 2011b. "A Word of the Empirics: The Ancient Concept of Observation and Its Recovery in Early Modern Medicine." *Annals of Science* 68 (1): 1–25. https://doi.org/10.1080/00033790.2010.495039

Potter, Jonathan. 1996. *Representing Reality: Discourse, Rhetoric and Social Construction*. London: Sage. https://doi.org/10.4135/9781446222119

Reinarz, Jonathan, and Rebecca Wynter (eds). 2015. *Complaints, Controversies and Grievances in Medicine: Historical and Social Science Perspectives*. London: Routledge.

Reisigl, Martin. 2017. "The Discourse-Historical Approach." In *The Routledge Handbook of Critical Discourse Studies*, ed. by John Flowerdew and John E. Richardson, 44–59. London: Routledge. https://doi.org/10.4324/9781315739342-4

Reisigl, Martin, and Ruth Wodak. 2001. *Discourse and Discrimination: Rhetorics of Racism and Antisemitism*. London: Routledge.

Reisigl, Martin, and Ruth Wodak. 2016. "The Discourse-Historical Approach (DHA)." In *Methods of Critical Discourse Studies*, ed. by Ruth Wodak and Michael Meyer, 23–61. 3rd edition. Los Angeles: Sage.

Richards, Jennifer. 2015. "Reading and Hearing *The Womans Booke* in Early Modern England." *Bulletin of the History of Medicine* 89 (3): 434–462. https://doi.org/10.1353/bhm.2015.0081

Simpson, Paul. 1993. *Language, Ideology and Point of View*. London: Routledge. https://doi.org/10.4324/9780203312612

Sommers, Sheena. 2011. "Transcending the Sexed Body: Reason, Sympathy, and 'Thinking Machines' in the Debates over Early English Midwifery." In *The Female Body in Medicine and Literature*, ed. by Andrew Mangham, and Greta Depledge, 89–106. Liverpool: Liverpool University Press. https://doi.org/10.2307/j.ctt5vj9mg.10

Staub, Susan C. 2011. "Surveilling the Secrets of the Female Body: The Contest for Reproductive Authority in the Popular Press of the Seventeenth Century." In *The Female Body in Medicine and Literature*, ed. by Andrew Mangham, and Greta Depledge, 51–68. Liverpool: Liverpool University Press. https://doi.org/10.2307/j.ctt5vj9mg.8

Taavitsainen, Irma. 2001. "Evidentiality and Scientific Thought-Styles: English Medical Writing in Late Middle English and Early Modern English." In *Modality in Specialized Texts: Selected Papers of the 1st CERLIS Conference*, ed. by Maurizio Gotti, and Marina Dossena, 21–52. Frankfurt: Peter Lang.

Taavitsainen, Irma. 2012. "Audience Guidance and Learned Medical Writing in Late Medieval English." In *Advances in Medical Discourse Analysis*, ed. by Maurizio Gotti, Marina Dossena, and Françoise Salager-Meyer, 431–456. Frankfurt: Peter Lang.

Taavitsainen, Irma, Päivi Pahta, Turo Hiltunen, Martti Mäkinen, Ville Marttila, Maura Ratia, Carla Suhr, and Jukka Tyrkkö. 2010. *Early Modern English Medical Texts: Corpus*. In *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen, and Päivi Pahta. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.160

Tatlock, Lynne. 2005. "Volume Editor's Introduction." In *The Court Midwife*, by Justine Siegemund, trans. and ed. by Lynne Tatlock, 1–30. Chicago: University of Chicago Press.

Thompson, John B. 1984. *Studies in the Theory of Ideology*. Cambridge: Polity Press. https://doi.org/10.1525/9780520312234

Walsh, Katharine Phelps. 2014. "Marketing Midwives in Seventeenth-Century London: A Re-Examination of Jane Sharp's *The Midwives Book*." *Gender & History* 26 (2): 223–241. https://doi.org/10.1111/1468-0424.12067

Whitt, Richard J. 2016a. "Evidentiality in Early Modern English Medical Treatises (1500–1700)." *Journal of Historical Sociolinguistics* 2 (2): 235–263. https://doi.org/10.1515/jhsl-2016-0014

Whitt, Richard J. 2016b. "Using Corpora to Track Changing Thought Styles: Evidentiality, Epistemology, and Early Modern Scientific Discourse in English and German." *Kalbotyra* 69: 265–291.

Whitt, Richard J. 2018. "'And all this is spoken of the naturall byrth …': Metadiscourse in *The Birth of Mankind* and Its German Source Text, *Rosengarten*." *English Text Construction* 11 (2): 225–255. https://doi.org/10.1075/etc.00010.whi

Wilson, Adrian. 1995. *The Making of Man-Midwifery: Childbirth in England, 1660–1770*. Cambridge, MA: Harvard University Press.

Appendix

Below is a table featuring the basic bibliographic information of the midwifery corpus compiled for the current project. Female works are indicated in **bold**, while an asterisk (*) placed by the author's name indicates no extant prefatory material.

Author	Title	Year of publication	Word count
1500–1599			
Eucharius Rösslin (Richard Jonas / Thomas Raynalde, trans.)	*The Byrth of Mankinde*	1540 (Jonas) / 1545 (Raynalde)	50,177
Jean Goeurot (Thomas Phayer, trans.)	*The Boke of Children* (part of *The Regiment of Life*)	1550	15,778
		Total 1500–1599:	*65,955*
1600–1699			
Edward Jorden	*A Briefe Discovrse of a Disease Called the Suffocation of the Mother*	1603	11,993
Jacques Guillemeau (unknown trans.)	*Child-birth or, The Happy Deliverie of Women*	1612	93,752
John Sadler	*The Sick Womans Priuate Looking-glasse*	1636	18,138
Jakob Rüff (unknown trans.)	*The Expert Midwife*	1637	55,269
Nicholas Culpeper	*A Directory for Midwives*	1651 (Part 1) / 1662 (Part 2)	69,308[16]
T. C., I. D., M. S., and T. B.	***The Compleate Midwife's Practice***	**1656**	**69,308**
Alessandro Massaria (unknown trans.)	De Morbis Foemineis, *The Woman's Counsellour*	1657	32,471
Jane Sharp	***The Midwives Book***	**1671**	**92,748**
François Mauriceau (Hugh Chamberlen, trans.)	*The Diseases of Women with Child, and in Child-Bed*	1683	121,817
Unknown	*Aristotle's Masterpiece*	1684	35,523
Robert Barret	*A Companion for Midwives*	1699	21,768
		Total 1600–1699:	*622,300*

---

16. A completely machine-readable version of Part 1 of Culpeper's text is not yet available, so only Part 2 is included here.

Author	Title	Year of publication	Word count
1700–1800			
Hendrik van Deventer (unknown trans.)	*The Art of Midwifery Improv'd*	1716	108,837
Pierre Dionis	*A General Treatise of Midwifery*	1719	116,680
Edmund Chapman	*An Essay on the Improvement of Midwifery*	1733	25,151
Sarah Stone	***A Complete Practice of Midwifery***	**1737**	**21,154**
Henry Bracken	*The Midwife's Companion*	1737	90,913
Fielding Ould	*A Treatise of Midwifery*	1742	42,934
William Clark	*The Province of Midwives in the Practice of their Art*	1751	8,562
Brudenell Exton	*A New and General System of Midwifery*	1751	22,812
George Counsell	*The Art of Midwifery*	1752	21,646
Benjamin Pugh	*A Treatise of Midwifery*	1754	30,100
John Memis	*The Midwife's Pocket-Companion*	1765	39,386
David Spence	*A System of Midwifery*	1784	99,798
John Grigg	*Advice to the Female Sex in General*	1789	71,154
*N. Torriano	*Compendium Obstetricii, or, A Small Tract on the Formation of the Foetus*	1753	15,507
Giles Watts	*Reflections on Slow and Painful Labours*	1755	20,030
Elizabeth Nihell	***A Treatise on the Art of Midwifery***	**1760**	**94,451**
William Smellie	*A Collection of Preternatural Cases and Observations in Midwifery* (Vol. 2)	1764	114,151
William Smellie	*A Collection of Cases and Observations in Midwifery* (Vol. 3)	1764	89,477
John Harvie	*Practical Directions, Shewing a Method of Preserving the Perinaeum in Birth*	1767	5,657
Robert Wallace Johnson	*Some Friendly Cautions to the Heads of Families*	1767	11,667
John Gibson	*Some Useful Hints and Friendly Admonitions to Young Surgeons on the Practice of Midwifery*	1772	4,135
Edward Foster	*The Principles and Practice of Midwifery*	1781	54,949
William Dease	*Observations in Midwifery*	1783	34,647

Author	Title	Year of publication	Word count
John Aitken	*Principles of Midwifery, or Puerperal Medicine*	1784	23,360
Stephen Freeman	*The Ladies' Friend*	1787	119,652
Alexander Hamilton	*Outlines of the Theory and Practice of Midwifery*	1787	73,538
Thomas Denman	*An Introduction to the Practice of Midwifery* (2 Vols)	1794–1795	85,447 (Vol. 1) + 111,124 (Vol. 2) = 196,571
Robert Bland	*Observations on Human and Comparative Parturition*	1794	40,692
Margaret Stephen	***Domestic Midwife; or, the Best Means of Preventing Danger in Child-Birth Considered***	**1795**	**65,212**
Martha Mears	***The Pupil of Nature; or Candid Advice to the Fair Sex***	**1797**	**44,876**
William Nisbet	*The Clinical Guide*	1800	82,772
		Total 1700–1800:	*1,790,471*

Unhappy patients and *eminent physicians*

The representation of patients and practitioners in Late Modern English medical writing

Anu Lehto

University of Helsinki

This chapter compares the representations of patients and doctors in eighteenth-century English in the corpus of *Late Modern English Medical Texts* (LMEMT). The representations are surveyed by collocation analysis in order to reveal shared beliefs linked with the social groups. Rather than individual collocates, the focus is on semantic preference and semantic prosody, which enable the analysis of abstract meanings and discourse-pragmatic functions of terms. Results indicate that physicians and surgeons are typically associated with education and professional expertise, and their social identity generally represents positive semantic prosody. By contrast, apothecaries are identified more negatively, displaying their lower status within the medical profession. Patients are often viewed as mere objects of treatment and their physical state is emphasised. However, patients receive sympathy as well, reflecting the philanthropic spirit of the era.

Keywords: representation, collocation, semantic prosody, semantic preference, patient, surgeon, physician, apothecary

1. Introduction

Medical texts of the eighteenth century feature a variety of subgenres, linguistic features and scientific developments, but at the heart of medical practice remain two essential participants that epitomise the purpose of medical practice, namely patients and practitioners. Although there are some previous studies on these groups in English historical medical texts (e.g. Lehto 2019a, Ratia 2020, Porter 1985), the social roles of patients and doctors in eighteenth-century medicine have not been earlier surveyed by corpus linguistic methods. This chapter, thus, uses collocation analysis to study how medical authors describe patients in comparison to three groups of practitioners: physicians, surgeons and apothecaries. The data comes from

the corpus of *Late Modern English Medical Texts* (LMEMT), which covers medical writing from 1700 to 1800 (see Taavitsainen and Hiltunen eds. 2019). The purpose is to identify common beliefs associated with the four social groups in medical discourse, and therefore the patients and practitioners are studied from the point of view of representations. Representations are defined as shared assumptions that prevail in texts or discourse communities (Baker 2014, Burr 2002). The characterisations in LMEMT can be expected to reflect widely held ideas on the social groups in eighteen-century medicine, as the corpus contains numerous subcategories of medical writing and provides texts written by both professional and lay authors for various audiences. However, the data emphasises the points of view of the medical writers, since the texts were not generally written by the patients themselves.

The representations are examined in the sociohistorical contexts of medical writing and practising medicine. Medicine experienced significant changes in the late modern period, as scientific experiments became common in acquiring more objective information on diseases (Risse 1992: 155). Philanthropy and polite correspondence also prevailed in the era. These changes can be expected to have affected the portrayals of patients and practitioners in the period. The representations are examined by collocation analysis, as it reveals more subtle meanings associated with the social groups (e.g. McEnery and Hardie 2012: 122–123). The collocates are analysed especially in terms of semantic preference and semantic prosody (Sinclair 2004: 31) that offer a more comprehensive idea of the themes and attitudes surrounding the headwords. This chapter argues that there were two opposite views on patients arising in the eighteenth century. On the one hand, the new medical methods and scientific thinking – with statistics, generalisations and experiments – detach the author from the single patient and shift the attention to groups of patients and generalisations. On the other hand, the authors emphasise a compassionate view towards patients' experiences especially in works aimed at lay audiences. This study further shows that the representation of patients differs from that of professionals, as physicians, for instance, are generally regarded as highly competent practitioners, giving instructions to patients.

2.   Medicine, practitioners and patients in eighteenth-century England

Eighteenth-century England was characterised by the Enlightenment, which affected ideas of society and scientific thinking. The era is often dated from the late seventeenth century up to the end of the eighteenth century. Enlightenment thinkers believed that there was a general orderly mechanism to be discovered and that this mechanism could explain the phenomena of the natural world. Mathematical methods gained ground especially in the first part of the eighteenth century and

reasoning was an important tool of explanation in natural philosophy (Reill 2003: 25). Isaac Newton's (1642–1727) experimental methods and search for fixed laws became influential. Consequently, medical authors explained functions of the body and diseases by systems of general laws to replace the dominant ideas of humoral theory and scholasticism of the previous centuries (Risse 1992: 155–156). Despite these developments, the eighteenth century is generally seen as an era in science that merely finalised the scientific revolution of the years 1500–1700, i.e. the period is considered to show only some early developments towards more modern medicine of the nineteenth century (Reill 2003: 23).

There were, nevertheless, numerous new improvements in eighteenth-century medicine. Especially towards the end of the century, medical authors relied increasingly on facts and on mathematical probabilities gained through observation instead of introspective reasoning (Reill 2003: 30–33, see also Taavitsainen, Jones and Hiltunen 2019). An important new method in the period was smallpox inoculation, which was "the first extensive immunization programme in history" (Risse 1992: 189).

Another novel theme was a growing interest in the promotion of general welfare and philanthropy in society. The upper classes contributed to the establishment of medical institutions, and many voluntary hospitals and dispensaries were first established in larger cities and later in the provinces, where the poor were given treatment for free (Porter 1995: 30–31). Polite manners and interaction were also noticeable features of eighteenth-century communication. Politeness shaped science and medical writing, as the upper classes became interested in scientific issues and participated in lay scientific discussions in coffee houses, for instance, and sent letters on their own experiences as patients to *The Gentleman's Magazine* (Taavitsainen 2019, Klein 2002: 890). Politeness cannot be solely related to the upper classes, but it can be seen as a general manner of communication in the era that facilitated interaction (Klein 2002: 873).

A strict hierarchy of medical practitioners continued to exist in the eighteenth century, with physicians being held in the highest esteem, followed by surgeons and apothecaries. Physicians practicing in cities and towns were usually university graduates, while surgeons would lack formal education and learn the profession by apprenticeship (Porter 1995: 29). Porter (1995: 29; 1992: 92) notes that the hierarchy between the professions did not truly reflect the reality of practicing medicine in the era: in 1704, London apothecaries were given the right to prescribe medicines and offer medical consultation, as opposed to merely distributing medicines. In addition, medical professionals in the provinces acted more as general practitioners representing all three professions. The status of medical practitioners advanced somewhat from the mid-eighteenth century onwards thanks to better education available in teaching hospitals (Loudon 1992). The newly established institutions

advanced the professionalisation of the occupations, as the hospitals – along with army and navy – provided new places of practice for physicians and surgeons. The Royal College of Physicians, already established in 1518, continued to promote the hierarchy and stricter regulation of the professions, allowing only university-graduated physicians to apply for membership in the society. However, the position of the practitioners and especially of the apothecaries did not truly improve until in the nineteenth century when new medical societies were established along with novel legislation, and when progress towards more modern medicine was made.

Patients from different social classes had varying access to medical treatment in the era. In general, patients were often more affluent people of the social ranks higher than the physicians themselves (Porter 1995: 24). In the eighteenth century, medical treatment became accessible to more patients from the middle and lower classes, as more people were able to pay for consultation, and the voluntary hospitals and dispensaries offered treatment for the poor (Risse 1992: 171). At the same time, the Enlightenment promoted more optimistic views on medicine urging patients to use medical services. That is, the medical authors as well as the upper and middle classes of society believed that all diseases were curable, as "[t]he aim of therapeutic measures was to aid and supplement the healing power of nature inherent in all human organisms" (Risse 1992: 150–152, 157). The physicians often took advantage of these ideas and the Newtonian general laws to explain the causes of diseases and functions of the body to their upper-class patients, who were familiar with these concepts. Although more numerous patients were able to seek healthcare in the era, it is worth noting that the household was still the most important arena for patients, as family members usually treated the sick at home (Withey 2019: 115–116).

3.   Evaluating representations by collocation analysis

Collocation in corpus linguistics refers to the tendency of words to appear near a central word or node. Collocation is thus based on frequency, as the purpose is to exclude arbitrary cooccurrences of lexical items (e.g. Stubbs 2003: 216; for methods on collocation see Gries 2013, Xiao 2015). In addition to frequency, collocations are specified by distance from the headword, which usually varies from one to four or five words (Sinclair 1991: 115). Brezina, McEnery and Wattam (2015: 140) state that a longer span is useful in analyses "interested in more general associations", and a span of around four words is usually considered suitable in capturing significant collocates affecting the meaning of the node (Sinclair et al. 2004: 13, see also Sinclair 1991, McEnery and Hardie 2012: 129). The span of five words is chosen for the present analysis, since the purpose is to focus on more general evaluations of the social groups, and this spread has been used in many previous studies on representation as well (e.g. Gabrielatos and Baker 2008).

Word meanings are affected not only by individual nearby words but also by groups of lexical items with similar meanings that can reveal speaker evaluation (e.g. Louw 1993: 159). Sinclair (2004: 31–33) draws attention to semantic preference and semantic prosody that distinguish more abstract meanings. Semantic preference is an association formed by groups of collocates that are semantically related (e.g. the words *see* and *visible* both refer to 'seeing' and collocate with *naked eye*). In contrast, semantic prosody indicates speakers' attitudinal meaning towards the node word or concept. Semantic prosody is hence "spread over a unit of language which potentially goes well beyond the single orthographic word and is much less evident to the naked eye" (Partington 2004: 131–132). For instance, Stubbs (2001: 45) shows that the verb *cause* conveys unpleasant ideas in phrases such as *cause problems*. Sinclair (2004: 34) summarises that semantic prosody is evaluated more on the pragmatic than semantic level and that it reveals the discourse function of words. Similarly, Hunston (2007: 258) underlines that "semantic prosody is a discourse function of a sequence rather than a property of a word". This principle can be applied to larger units as well, since specific genres tend to prefer certain collocates and activate dissimilar attitudinal meanings and semantic prosodies (Hunston 2007: 263). Therefore, the medical texts likely portray associations that differ from other fields, and consequently the findings made on the portrayals of practitioners and patients cannot be straightforwardly generalised to other genres, such as legal texts or patients' diaries.

The concept of semantic prosody resembles the idea of representations of social groups. Representations are shared and usually unrecognised assumptions about groups of people or concepts in a discourse community (Burr 2002: 106). Since representations are collective beliefs, they do not convey personal attitudes of individual speakers (Stubbs 2001: 215). These common associations are formed by repetition, which also strengthens the ideas in the community. The analysis of semantic preference and semantic prosody aims at tracing these implicit and repeated cultural assumptions. As collocations are not diachronically static, semantic prosodies can change in the course of time. Static collocates, which remain in use for long periods of time, nevertheless, give a more accurate portrayal of the headword if compared to collocates that emerge only at certain points in time, e.g. in newspapers, debates on a particular topic can prevail in the news for a very limited time span (Gabrielatos and Baker 2008: 11; see also McEnery and Baker 2019). The present study does not divide the material into separate time periods, but evaluates the possible diachronic changes to semantic preference and semantic prosody qualitatively.

There are some earlier works that explore representations in historical genres, and a few of them rely on collocation analysis. Ratia (2020) analyses the ten strongest collocates for *patient(s)* and states that in the early modern era the patient is shown as the recipient of treatments, while in the eighteenth century many collocates

emphasise patients' experiences (e.g. *feels*). Furthermore, Lehto (2019a) indicates that the wellbeing of patients is dealt with in texts on PUBLIC HEALTH in LMEMT and that new themes, such as statistical methods and observations, emerge. Patients' experiences have been recorded in works by historians as well. Porter (1985) studies patients in *The Gentleman's Magazine*, and Lane (1985) explores patients' narratives in eighteenth-century letters and diaries. Lane (1985: 216–217) finds that patients, for instance, complain on treatments they receive but also express thankfulness for overcoming their diseases. McEnery and Baker (2019) – as an example of studies on representation in other historical genres – show that different collocates are linked to the criminalised poor, such as beggars and rogues, affecting their portrayal in seventeenth-century texts (see also Lehto 2019b). In addition, a number of studies explore representation of genders in historical and present-day texts and study social groups in newspapers (e.g. Bäcklund 2006, Gupta 2015, Sveen 2010, Baker and McEnery 2005, Pearce 2008).

4.   Data and method

4.1   The Corpus of *Late Modern English Medical Texts*

The *Late Modern English Medical Texts* corpus (LMEMT, Taavitsainen et al. 2019) is a register-specific corpus comprising texts from different fields of eighteenth-century medicine in England (see Taavitsainen and Hiltunen 2019, see also Hiltunen, this volume). The corpus covers one hundred years of medical writing from 1700 to 1800 and contains about two million words. LMEMT comprises the following six main topical categories as well as two additional categories that relate to publication format, i.e. texts from scientific and general periodicals.

Main text categories

1. GENERAL TREATISES
2. SPECIFIC TREATISES
    A.   DISEASES
    B.   METHODS
    C.   THERAPEUTIC SUBSTANCES
    D.   MIDWIFERY
3. MEDICAL RECIPE COLLECTIONS
4. REGIMENS
5. SURGICAL AND ANATOMICAL TEXTS
6. PUBLIC HEALTH

SCIENTIFIC PERIODICALS
- *Philosophical Transactions*
- *Edinburgh Medical Journal*

GENERAL PERIODICAL
- *The Gentleman's Magazine*

The present study is carried out in the six main categories of LMEMT. The three periodicals are excluded, as they represent different publishing formats. The texts in LMEMT vary in terms of their audience and authors: most texts are written by university-educated medical professionals addressing professional readers, but there are also many texts targeted at lay readers and works written by authors lacking formal education. The intended audience – usually mentioned on the title page or preface – affects the choice of linguistic features and style used by the author (see Bell 2001).

4.2 Methods

The first step in the analysis was to generate a word list from the corpus to examine which words are used to refer to patients and medical practitioners. The word list and collocates were identified by using the AntConc software (Anthony 2014). The nodes *patient, physician, surgeon* and *apothecary* were chosen, since they are frequent in the data and represent the different types of practitioners this study is interested in. The list of nodes is by no means exhaustive, e.g. the word *dr.* refers to practitioners but it is solely used as an address term and precedes names of specific individuals.[1] The nodes include different word forms in addition to the lemmas (e.g. *patient, patient's* and *patients*) and possible spelling variants (*surgeon* and *chirurgeon*). Spelling variants are treated as separate words by corpus software, but this characteristic affects only few collocates in the data, since spelling is already rather consistent in the eighteenth-century texts under investigation.

The second phase in the study was to set the span and frequency thresholds for the collocates, and the span of five words and a minimum frequency of four occurrences were chosen. The threshold corresponds to the normalised frequency of 3.1 in million words, although most collocates show higher frequencies, ranging from 3.1 (4 instances) to 100.2 (131 instances). The mutual information (MI) scores of the collocates vary from 1.3 to 12.2, but only one collocate has an MI score below 2 and altogether 32 collocates show scores below 3.[2] Since studying each category

---

1. Among other practitioners, there are instances of *druggist*(*s*) (22 occurrences), *nurse*(*s*) (122), *practitioner*(*s*) (160), *gentleman/-men* (293), *midwife* (106) and *doctor* (252, e.g. *doctor of physic*).

2. According to Hunston (2002: 71), collocates that show MI scores of 3 or above are significant collocates.

separately did not generate enough collocates for all headwords for systematic categorisation of the topics, the collocates were retrieved from the six main corpus categories.

The third step in the analysis was to manually exclude grammatical collocates including conjunctions and articles from the generated list. After that, the content collocate types and tokens were studied individually in order to determine their meaning in context, i.e. each concordance line was explored and numerous tokens were investigated in their wider textual context. Each collocate type was then categorised according to their meaning in order to evaluate the semantic preference of the nodes, and 16 themes were recognised.

5.   Distribution of headwords and semantic categorisation of collocates

The headwords *patient, physician, surgeon* and *apothecary* occur 4,149 times in the six analysed categories in LMEMT. The word *patient* is the most frequent headword with 1,724 instances, followed by *physician* (1,344 instances), while *surgeon* (629) and *apothecary* (452) are noticeably less frequent. Altogether 1,448 collocate types and 32,123 collocate tokens were found near the headwords. As the study excludes grammatical collocates, the number of analysed content collocate types is 1,052. The distribution of content collocate types reflects the frequency of headwords, *patient* attracting the highest number of collocate types (451) followed by *physician* (335), *surgeon* (165) and *apothecary* (101), as is indicated in Figure 1.
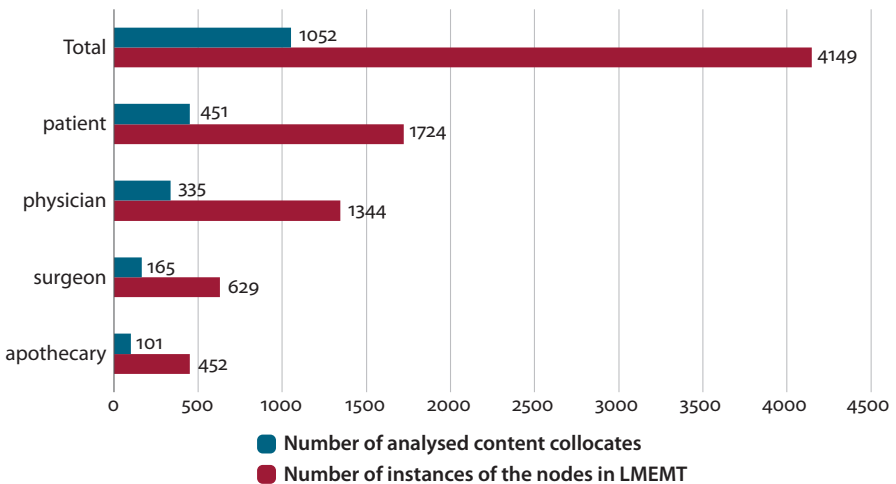
Figure 1. Raw frequencies of nodes and content collocate types in six categories of LMEMT[3]

3.   All of the corpus searches were carried out in a pre-release version of the LMEMT corpus.

The distribution of headwords varies noticeably among the corpus categories of LMEMT. It shows the varying contexts in which patients and practitioners are considered, affecting the choice of accompanying words and associations, e.g. texts on PUBLIC HEALTH specifically address the practitioners and patients from the point of view of hospitals and society, while recipes offer more practical advice for treatment. The word *patient* occurs most often in the categories of SURGICAL AND ANATOMICAL TEXTS, PUBLIC HEALTH and in the subcategory of DISEASES in SPECIFIC TREATISES. The node *physician* is most frequent in PUBLIC HEALTH and GENERAL TREATISES, and *surgeon* is most often mentioned in PUBLIC HEALTH and SURGICAL AND ANATOMICAL TEXTS, while *apothecary* appears mostly in PUBLIC HEALTH, MEDICAL RECIPE COLLECTIONS and in the subcategory of THERAPEUTIC SUBSTANCES. The normalised frequencies of the headwords in 100,000 words within each corpus category can be seen in Figure 2.
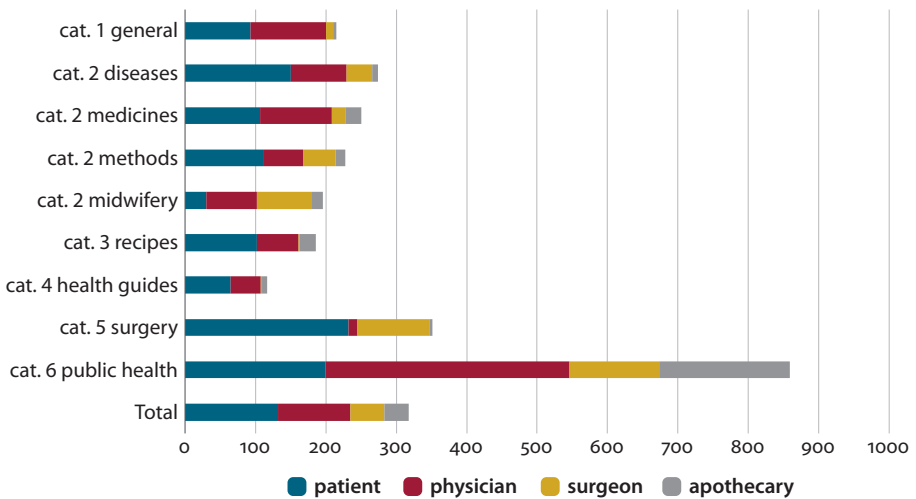
Figure 2. Normalised frequencies of the headwords by categories in LMEMT (f./100,000 words)

The authors mention patients and practitioners well over twice as often in PUBLIC HEALTH than in SURGICAL AND ANATOMICAL TEXTS, and the difference is sevenfold if compared to REGIMENS.

In order to evaluate the semantic preference linked to the social groups, the collocates were categorised into 16 classes by their topic. The semantic categories include collocates on actions (e.g. verbs such as *go* and nouns including *admission*), abstract concepts and objects (Topics 1–3 below). There are also references to medical professionals (e.g. *nurse*) and descriptive references to medical practitioners (*skilful*) as well as references to patients (*young*). Additionally, the collocates encompass instances of other members of society (*lord*) and some references to the

ancient authorities (e.g. *Hippocrates*) (Topics 4–8). Separate classes were established for medical topics, that is, diseases, body parts and bodily functions, symptoms and medicines (9–12). Some topics, like location, date and time as well as quantities (13–15), are more general. Finally, the last category encompasses mixed collocates, i.e. the tokens fit into more than one category (e.g. *part* in the sense of 'body parts' and 'three parts'). The topics are listed in Table 1.

Table 1. Topics of content collocates (semantic preference)

Topic	Topic name	Example(s) of collocates
1.	Actions	*been, have, advise, admission*
2.	Abstract concepts	*knowledge, medicine*
3.	Objects	*letter, knife*
4.	Reference to medical practitioners	*nurse*
5.	Descriptive references to medical professionals	*eminent*
6.	Reference to other persons	*lord*
7.	Reference to the ancient authorities	*Hippocrates*
8.	General reference to patient	*robust*
9.	Diseases	*smallpox*
10.	Body parts and bodily functions	*head, respiration*
11.	Symptoms of diseases and condition of the patient	*tender, weakening*
12.	Medicines and cures	*diet, mercury*
13.	Location	*London*
14.	Date and time	*year, tomorrow*
15.	Quantities	*five, several*
16.	Mixed tokens	*part*

The category of **actions** (Topic 1) encompasses only actions which are not related to the other semantic classes. Actions referring to functions of the body, for instance, are listed under **body parts and bodily functions** (Topic 10).

6.   Semantic categorisation of collocates for *patient*

The collocates for patient spread over 14 of the 16 themes, as there are no references to authorities or descriptive references to medical professionals.[4] The topic of **actions** is by far the most common, followed by **symptoms and condition of the patient, quantities, medicines and cures** and **abstract** concepts. Figure 3 shows the distribution of the themes in percentages.

---

4.   The strongest collocates are *affluent, diabetic, complains, payments, expired, maintenance, attendance, coughed, suffocated, aversion, convenience, administered, asylum, visited* and *sit*. The raw frequencies of these collocates range from 4 to 7 instances and the mutual information (MI) score varies from 7.2 to 9.3.
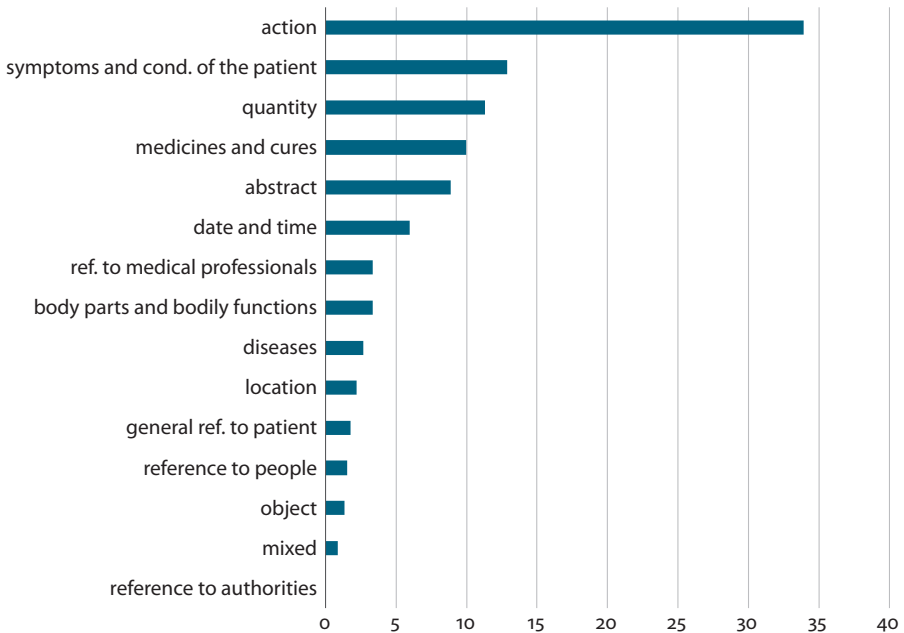
Figure 3. Semantic preference for *patient* in six categories of LMEMT (percentages)

Actions are hence the most common theme surrounding patients, encompassing 33.9% of the collocates. This topic is frequent, as it combines a vast array of general actions and also stance expressions and mental verbs. Most of the collocates are expressions relating to the daily practice of visiting and caring for patients (*visit, attend, practise, recommended* and *advice*). The following concordance lines (Examples 1–3) illustrate these common actions of prescribing medicines and treating patients in the categories of MEDICAL RECIPE COLLECTIONS, THERAPEUTIC SUBSTANCES and SURGICAL AND ANATOMICAL TEXTS. The collocates referring to actions are marked in bold and some illustrative examples of word strings that include collocates on other topics are italicised (emphasis in all examples mine):

(1)   Then strain it out, and ***let*** the *Patient **take*** four ounces at a time, *twice a Day*.
      (MEDICAL RECIPE COLLECTIONS, Kettilby, *Receipts*, 1714: 180)

(2)   In this mode of prescribing, when I had so many *patients* to ***attend*** to in the *space of one, or at most of two hours* […]
      (THERAPEUTIC SUBSTANCES, Withering, *Foxglove*, 1785: 3)

(3)   The general preparation is that which is ***recommended*** to every *patient* afflicted with the Cataract […] First, ***take*** from the *patient* his neckcloth, unbutton his shirt collar, and […]   (SURGICAL AND ANATOMICAL TEXTS, Bischoff, *Extraction of the Cataract*, 1793: 11, 34)

As shown in Examples (1) and (2), the instructions for taking medicines and performing cures are often accompanied by collocates relating to **quantities** and **date and time**, e.g. *four ounces* […] *twice a Day*. Example (3) is from SURGICAL AND ANATOMICAL TEXTS, and this as well as many of these works give instructions on preparing and positioning patients for surgery. In all these contexts, patients are portrayed as objects of the medical procedures and cures, and they are discussed in relation to medicines and other treatments. Lane (1985: 228) similarly mentions that in eighteenth-century patients' letters and diaries "[m]ost medical advice was authoritarian, and the practitioner usually 'declared', 'pronounced' or was 'of the opinion'" and "[e]ven the most eminent patient was expected to be obedient to medical instructions". Although the expressions *declared* and *pronounced* are not found as collocates for patients, the idea of the physician having an authoritarian status is regularly visible in LMEMT when the patient is subjected to treatment.

Within the category of **actions**, stance expressions also appear in contexts where the authors give instructions to patients and practitioners treating them. Stance expressions add up to 7.8% of actions, and especially the modals *may, must, will* and *should* are frequent in passages such as *The Patient must neither eat nor drink any thing* (Anonymous, *Pilulae Polychrestae*, 1719: 11) and *patient may take a dose of rhubarb* (Cole, *Ladys Complete Guide*, 1788: 520). Furthermore, 9.2% of collocates categorised as actions are mental verbs, e.g. the verbs *think* and *know* are used to evaluate the effects of cures and to consider varying views on diseases (e.g. *we think the patient is in a state* and *as he may think best for his patient*).

The collocates on **actions** further show that patients are viewed through observations and experiments. In the subcategory of DISEASES, Morley, for instance, verifies his views on curing tumours and ulcers by narrating various Observations and case studies:

(4)   […] for *in my Course of Practice I have had this verified by repeated **Observations***: I have had a *Patient* afflicted with Tumours and Ulcers from the Crown of the Head to the Sole of the Foot, and in above twenty different Places besides at once, the major Part of which have been cured by the Application of dry Lint only […]   (DISEASES, Morley, *King's Evil*, 1763: 21–22)

Morley lists several case studies of patients and then proceeds to general conclusions that indicate the typical progression of illnesses and the effects of the cure. The medical authors often rely on experiments and make generalisations on the basis of these observations, and instead of focusing on single cases, the patients are categorised into similar groups.

The topic of **symptoms and patients' condition** is the second most common category of the collocates (12.9%). The collocates include *coughed, suffocated, pain, heat, strength, fit, sweat* and *incurable*, and the patient's state is often described as

weak, quiet and *dying* (e.g. *Where the Patient is Weak, and very Restless, 'tis best to mix a third part Venice Treacle* in Kettilby's *Receipts*, 1714: 177). The collocates *complain, complaint* and *felt* are used in indirect quotations to record symptoms mentioned by patients: *patient complains of pain in swallowing* and *I carefully write down every Patient's Complaints*. It is noticeable that especially the collocates on symptoms and condition of the patient often show compassion towards the patients. The collocates *poor, suffer, unhappy, relief, miserable* and *unfortunate*, for instance, occur when the condition of the patient is evaluated. This sympathetic attitude is particularly noticeable in texts aimed at lay audiences:

(5)   *What a moving Spectacle* of mortal Nature is it, to see the **unhappy** *Patient* extended on the Rack, *groaning and crying* out in *Agonies of Distress and unspeakable Torment*: […] Distracted *with his Sufferings*, he lies wakeful, count-ing the Hours, any one of which, when protracted and multiplied by *raging Misery*, seems a numerous Train, so sluggish and unprogressive does Time grow under *grievous Sufferings*. […] they [the agonies] soon recover their Strength, rekindle their *Rage*, and *insult the Patient* with as great *Fury* as before. […] I say, the Stone, […] attended with a *terrible Train of Symptoms*, that *render the Patient a great Object of Compassion* […]

(DISEASES, Anonymous, *Gravel and Stone*, 1738: 2–3)

The anonymous author writes about patients suffering from kidney stones and re-fers to a *moving spectacle, unhappy Patient, unspeakable Torment* and *raging Misery*. At the end of Example (5), the author explicitly alludes to compassion: *render the Patient a great Object of Compassion*. Sympathy is also expressed in texts that con-cern the controversy between physicians and apothecaries. The authors who oppose apothecaries practicing medicine portray patients as *unhappy* or *miserable patients*. This is done, for example, by Lucas (*Pharmacomastix*, 1741: 33) who states that *Notwithstanding all this gross abuse of the miserable patient, this unworthy apoth-ecary* […] *makes a most extravagant large bill*. He repeatedly refers to patients by the word *sufferer*, which is found in texts on hospitals as well, as these works aim to appeal to the subscribers. The purpose of expressing sympathy in these contexts is thus to strengthen the idea that patients suffer from treatments received from uneducated apothecaries.

The topic of **quantities** is the third most frequent theme, accounting for 11.3%. These collocates include *one* (as in *one of these patients*), *many, plentifully* and *small*. Quantities are used to give detailed instructions on medicines together with the collocates on date and time. The collocates on quantities indicate that patients are viewed through numbers. This practice is strong in texts that consider the role of hospitals in the category of PUBLIC HEALTH, e.g. the collocate *twelve* in *the General Dispensary will annually afford relief to about twelve thousand patients* (Lettsom,

Improvement of Medicine in London, 1775: 17). In hospital texts, patients are regularly related to money and costs, e.g. *payments* (and *number*) is a strong collocate:

> (6)   […] the [York Lunatic] Asylum contains *Sixtysix patients*, and these are formed into *three divisions*. The first division (in *number twenty*, […] consists of such *patients* as are *of better condition*, who *pay* a considerable *weekly sum* for their board and medicines, the *Surplus* of which is employed in lessening the **payments** of those *patients* who […] The Second division (in number *twentysix*, of which upwards of *one fourth are parish paupers*), consists of those who *pay eight Shillings* per week […]; a *Sum* which from experience is found equal to the *expence* […] all the advantages arising from the surplus of the payments made by the *patients in affluent* circumstances […] The third division (in **number twenty**) consists of *patients in low* circumstances who *pay* […]
>
> (PUBLIC HEALTH, Hunter, *Letter to York Lunatic Asylum*, 1788: 8)

Hunter introduces the York Lunatic Asylum, explaining the types of admitted patients and the billing system. These themes along with the strong collocate *asylum* in Example (6) display the growing importance of hospitals that shifted the focus from individual patients to classes of similar cases and increasingly associated patients with numbers. In addition, some collocates on quantities (e.g. *eight*) reveal that patients are considered from the point of view of statistical information at the end of the century. In his *Inquiry into the Cow Pox* (1798: 26), Jenner calculates that among inoculated patients *eight patients presented themselves who had at different periods of their lives had the Cow Pox* but none of them *received the variolous infection*. Furthermore, Example (6) contains the word *circumstance*, which is an example of **abstract concepts** that cover 9.4% of the collocates. Other examples of abstract collocates include *life* (e.g. *save the patient's life*), *health, nature, natural, security* and *account*. The concepts of life, death and health are central themes of medical procedures and the outcomes of the treatments are regularly stated (e.g. *the patients were restored to health*).

Collocates referring to **medicines and cures** are the fourth most common theme rising to 10.0%, and they are firmly related to the actions on examining patients and practicing medicine. Instead of specific cures or remedies, the repeated collocates are hypernyms: *medicines, cure, method, rest, drink, diet, regimen, wine, purging, exercice, diet* and *water* (in spas). The collocates often refer to a regimen, relating to the way of life of the upper classes, and they reflect the high number of medical works published for lay audiences in the eighteenth century (Porter 1992: 106). In Example (7), *regimen* and *proper* are examples of medicines and cures that collocate with *patient*:

(7)   A ***proper regimen***, with the *patient's* own *endeavours*, act mostly towards the
cure; […]
(GENERAL TREATISES, Robinson, *Every Patient his own Doctor*, 1785: 24–25)

Patients have a more active role in Example (7): they are portrayed as being responsible for their own wellbeing by following physicians' advice. The example reflects the optimistic view of the century and regimen texts in general, emphasizing that diseases could be combatted with suitable treatment. Within the collocates on medicines and cures, the only reference to a particular cure or method is *inoculated*, as the new procedure was introduced in the eighteenth century. This collocate appears mostly in texts for professional audiences (e.g. Jurin's *Inoculating the Smallpox*, 1724), but collocates are also found in the anonymous *Inoculation Made Easy* (1766), which instructs lay readers to inoculate smallpox by themselves. The text guides the reader through proper meals, preparative powders and notes possible side effects, indicating that *In about Six or Seven Days after the Inoculation, the Patient Sometimes feels a Kind of a Slight Pain in the Inoculated Arm* […] *but all those Symptoms Soon go off* (1766: 9).

7.   Semantic categorisation of collocates for *physicians, surgeons* and *apothecaries*

The collocates for *physicians, surgeons* and *apothecaries* most often focus on **actions, medical professionals** and **quantities** as well as make **descriptive references to the practitioners**.[5] Collocates referring to **actions** are the most common among each group of practitioners, but there is some variation in the distribution of other categories. For physicians, **descriptive references** rise to second place, as their personal characteristics are regularly highlighted, and **abstract concepts** are more frequently found than for surgeons and apothecaries. For all three groups of practitioners, references to **date and time**, and **medicines and cures** are rather important collocates, while **references to patients**, their **symptoms** and **body parts** are rare. The distribution of the topics in percentages is illustrated in Figure 4.

---

5.   The strongest collocates for physicians are *chirurgeons, prescripts, consults, chirurgeon, arabian, ablest, fees, consulting, eminent, educated, philosophers, antient, experienced* and *skillful*, showing frequencies from 4 to 32 and mutual information (MI) scores from 9 to 10.6. For surgeons, the strongest collocates are *junior, senior, elected, navy, Suffolk, appointment, limited, Robert, Jan, experts, appointments, incorporated, barber, chirurgical* and *eminent* with frequencies from 4 to 14 and MI scores from 9.6 to 13.2. For apothecaries, the strongest collocates are *wages, druggists, salary,* (*behalf*), *gains, shop, shops, dispensaries, practising, quacks, pharmacy, trade, qualified, doctors* and *medicins.* Their frequencies vary from 4 to 12 and MI scores from 8.7 to 11.5.
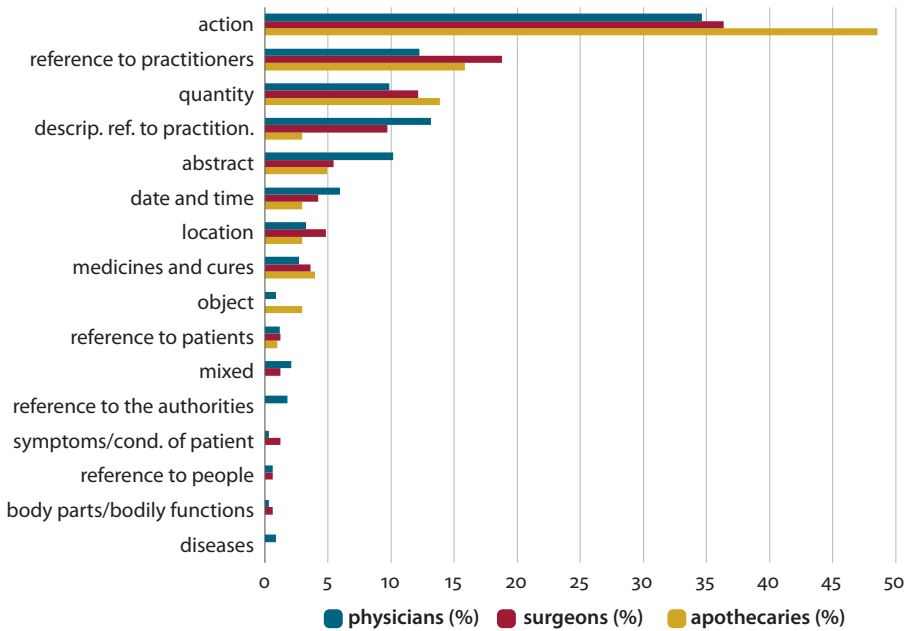
Figure 4. Semantic preference for medical practitioners in LMEMT (percentages)

The collocates on **actions** clearly form the most frequent and varied category of topics again, accumulating from 34.6% to 48.5% of the cases. Similarly to patients, the collocates include rather general actions indicated by verbs such as *take, make, use, give* and *said* and nouns such as *practice*. The category also comprises mental actions such as *know, think* and *suppose* (12.9% of actions related to physicians and 3.3% to surgeons) as well as expressions of stance, e.g. *must, may* and *shall* (9.5% of actions of physicians and 11.7% of surgeons). Thus, many authors refer overtly to mental actions or possible limitations in expertise:

(8)   has often since made me **think**, that physic and *physicians* are but imperfect instruments to restore to vigour, constitutions worn out with cares and anxiety […]   (DISEASES, Jones, *Nervous Fevers*, 1789: 25)

(9)   the remedy often proves worse than the disease: For it is *well **known*** to *physicians*, that all fevers which come after long purging […] are of the worst kind, and frequently prove fatal.
               (GENERAL TREATISES, Morland, *Causes of Chronic Diseases*, 1774: 29)

The collocate *know(n)* is common within all three professional groups, and it can among other pragmatic functions refer to the knowledge of practitioners and general knowledge of people. References to knowledge were especially frequent in

previous periods of medical writing in the corpus of *Middle English Medical Texts* (MEMT, 1375–1500, Taavitsainen et al. 2005) and *Early Modern English Medical Texts* (EMEMT, 1500–1700, Taavitsainen et al. 2010), although the overt references show a declining pattern diachronically (Hiltunen and Tyrkkö 2009: 76). The theme of knowledge is still associated with practitioners in LMEMT, and it emerges in the other semantic categories as well, e.g. the collocate *knowledge* is here classified in abstract concepts. Further, the verb *said/says* is often used to cite previous findings and arguments:

(10)   Objection I. ’Tis Demonstrated by the Mathematicians, that neither Corpuscles in the Aqueous or Crystalline Humor can be perceiv’d on the Retina; and ’tis ***said*** by *Physicians*, that a Beginning of a Cataract is a Suffusion of the Eye

<div align="right">(SURGICAL AND ANATOMICAL TEXTS,
Taylor, Mechanism of the Eye, 1727: 41)</div>

Taylor answers numerous objections presented in earlier works on cataract. The collocates for apothecaries reveal similar descriptions of general actions, but the proportion of stance expressions is higher, rising to 18.4% of actions, while mental actions comprise 4.1% of the cases. The authors regularly express their attitude about apothecaries, and in general, the collocates indicate more negative wordings such as *pretend* and *ignorance* as in *But let no Apothecary pretend to prescribe; his Business is only to obey Prescriptions* (Anonymous, *An Essay Reforming the Modern Way of Practicing Medicine in Edinburgh*, 1727: 21).

References to practitioners are an especially common collocate type for surgeons (18.8%) but also for apothecaries (15.8%) and physicians (12.2%). A reoccurring practice within the collocates for surgeons and physicians is to refer generally to other professionals (*practitioners, physicians, apothecaries, surgeons*, [*man*] *midwife* and *men*) or fields of study (*naturalists* and *chemists*). The word *college* ranks high among the strongest collocates, as the authors often associate surgeons and physicians with the *College of Physicians*. In addition, numerous **references to named professionals** are found. The collocates for physicians and particularly for surgeons include proper names such as *John* and *William* as well as general terms of address, including *dr.* and *sir*, e.g. "About ten years since, John Paul Baumer, a young physician of Franconia, undertook to improve this method" (Büchner, *Method to Enable Deaf Persons to Hear*, 1770: 33). The collocates for apothecaries include similar general **references to practitioners** (e.g. *druggists, surgeons* and *doctors*) and proper names and addresses. The collocate *quacks* is, however, a strong collocate for apothecaries, which expresses negative evaluation towards apothecaries treating the sick. In the following example, apothecaries are related to quacks and nurses and their practicing of medicine is questioned:

(11)   And as Nature alone will conquer Some Diseases, […] even the Mismanagement
of *Nurses, Apothecaries* and other **Quacks**, is not able to hinder the Recovery
of very large Numbers, which fall under their Hands.

(PUBLIC HEALTH, Guybon, *Growth of Empiricism*, 1712: 54–55)

The authors criticise apothecaries for lacking formal education and relate their skills
to nurses, accusing their practicing of medicine as *the most fatal evil*. If compared
to surgeons and physicians, the word *college* does not surface among the collocates,
underlining the apothecaries' less educated position in the field of medicine.

Descriptive references that evaluate medical practitioners are typical collocates
and define especially physicians, adding up to 13.1% for physicians and to 9.7% and
3.0% for surgeons and apothecaries, respectively. These collocates usually appear
as attributive adjectives within all medical groups. For physicians and surgeons,
the collocates emphasise their skills, education, character and reputation, e.g. *eminent, great, good, best, skilful, experienced, honest, reputation, qualified, educated*
and *excellent*. In Example (12), the author uses the attributive adjectives *prudent*
and *skilful*, and the whole excerpt compliments physicians on their hard work in
acquiring education:

(12)   […] that it is no easy Matter to become *a **prudent** and **skilful** Physician*: Much
Reading, much Thought and Pains will really be required; many Things are
necessary to be learnt much Diligence in Inquiry, […]

(GENERAL TREATISES, Rudd, *Method to know the disease*, 1742: 11–12)

Many of the descriptive collocates that define surgeons and physicians appear
among the strongest collocates and form frequently repeated word pairs (*great physician*), suggesting that they can be seen as formulaic sequences (see Wray 2002: 9).
Especially the wordings *eminent, great, skilful, good, educated* and *experienced physician/surgeon* reoccur throughout the eighteenth century. Example (13) illustrates
the repetitious use of *eminent physician(s)* in an anonymous work on the nervous
system from 1780:

(13)   Many **Eminent Physicians**, regarding all disorders of the Nerves […] would
effect a cure by means of remedies which are hot, and strengthening, by Gums,
[…] *which is an error similar* to the following, in speaking of indigestion and
Stomachic remedies. […]
That **eminent Physician** *Mr Lewis*, says, That of all Remedies, whether external
or internal, there is none can equal the virtue of the Cold Bath […]
I shall conclude this Section with Dr Chene's observation, in his English Malady,
the reading of which, strikes the Human Mind with astonishment: But of all
the miseries that afflicts human Life, (says that **Eminent Physician**) […]

(GENERAL TREATISES, Anonymous, *Dissertation upon the nervous system*,
1780: 11, 19, 28)

The word *physician* appears in the work five times and in four cases it is preceded by the collocating attributive adjective *eminent*. This positive representation is repeated even when the author mentions physicians' cures, which are not approved by the author. For surgeons, the collocates *junior* and *senior* (and *navy*) emerge as strong collocates that describe the type of surgeons in hospital texts, illustrating the role of new institutions for these practitioners.

The physicians and surgeons are thus surrounded by positive **descriptive references**, whereas descriptive collocates for apothecaries are not common and the references are negative. Charles Lucas reflects this view in his *Pharmacomastix* (1741: 7) by stating that *dishonest apothecaries […] may destroy the reputation of the most careful and judicious physicians and chirurgeons*. He also accuses apothecaries of murdering patients and calls them *blockheads* and *ideots*. Some individual texts, nevertheless, exhibit positive representations of apothecaries that do not surface as collocates, e.g. the adjectives *proper* and *skilful* (*apothecary*) appear in *Reasons* from 1704 in which the anonymous author named Apothecary defends the skills and education of the practitioners. There are some collocates in other texts that could be assumed to indicate positive wordings such as *worthy apothecary*, but even these phrases are used negatively in their textual contexts. Generally, the range of unflattering collocates for apothecaries shows that these practitioners are widely surrounded by negative associations, even in contexts where they are not criticised.

Abstract concepts are the fourth most common theme of the collocates for physicians (10.1%) and apothecaries (5.0%), while descriptive references rank fourth for surgeons (9.7%). **Abstract concepts** include collocates such as (*healing*) *art, physic, nature, knowledge, life, right* and *advantages*. The word *nature* collocates with physicians, as their treatments are compared to the healing power of nature: *'Tis the part of a Physician to assist Nature, by observing nicely the various steps she takes to repulse the Enemy* (Catherwood, *Curing the Apoplexy*, 1715: 16). The collocate (*healing*) *art* is discussed in relation to physicians and surgeons. Furthermore, the collocate *right* illustrates the growing urge to regulate medical professionals and the field of medicine by legislation. Christopher Stanger explores statutes and the charters of the College of Physicians and argues that also physicians educated abroad should be allowed to become licentiates in the College:

(14)   We Shall next proceed to prove, by decisive evidence, that the letter and Spirit of the **charter**, which originally incorporated the *physicians* of *London*, and of the act of parliament which confirmed it, clearly demonstrate, that the crown and the legislature did thereby grant and convey, in perpetuity, to every *competent physician, the* **right** of being admitted a member of the corporation, if found, upon examination, possessed of the qualifications required by the charter and act of parliament.   (PUBLIC HEALTH, Stanger, *Justification of the Right of Every Physician*, 1798: 39)

The topic of **quantities** ranges from 9.8% to 13.9% and comprises collocates such as *more, one, two, little* and *all*. The expressions of quantities that are found near the headword *patient* typically refer to the use of medicines and cures. In contrast, the collocates for the practitioners are usually not connected to the practicing of medicine but to numbers of persons or concepts:

(15)   There were *two* Physicians, *two* Surgeons, and an Apothecary, present at the Operation […]

            (SURGICAL AND ANATOMICAL TEXTS, Douglas, *Lithotomia*, 1720: 22)

(16)   there would be *little* Use for *Physick* or *Physicians* in chronical *Cases* […]
            (REGIMENS, Cheyne, *Essay of Health and Long Life*, 1724: 38)

The proportions of collocates related to **medicines and cures** vary from 2.7% to 4.0%. The collocates connect surgeons with operations (*operation, care* and *opened*), e.g. *The trachea, when opened by Mr Wood surgeon, was covered* and display physicians performing cures (*prescriptions, medicines, care, cure, remedies* and *application*). Apothecaries are associated only with the collocate *medicine(s)*. The topic of **diseases** relates only physicians to *fevers* and to the general collocate *diseases*; the collocates thus comprise general terms instead of specific cures or illnesses. Collocates defining **locations** connect physicians, for instance, with *London, dispensary, room, city* and *hospital*, and similarly surgeons are linked to *London* and *hospitals*, while for apothecaries the collocates include *shop*(s) and *dispensaries*. **References to the ancient authorities** are rare, as only six collocates emerge for physicians (e.g. *Hippocrates*). However, the collocates *antient* (*physicians*) and *philosophers* (and *Hippocrates*) rank among the strongest collocates for physicians, as contemporary knowledge and medical practices are regularly compared to previous works.

8.   Discussion and conclusion

The studied collocates indicate that patients and practitioners are surrounded by similar themes in eighteenth-century medical writing, but specific topics are emphasised for each social group and result in dissimilar semantic prosodies. The topics show that the patients are mostly linked to medical procedures and medicines and their symptoms are carefully recorded. In addition to treating illnesses, patients are associated with the method of observation. The focus shifts increasingly from one patient to similar groups of patients and the authors draw general conclusions on diseases on the basis of experiments. The hospital movement that characterised the eighteenth century promotes the tendency to classify patients with similar

illnesses together, and the texts especially in the category of PUBLIC HEALTH relate hospital patients with financial issues including their expenses in these institutions. Furthermore, numbers regularly collocate with patients, as quantities are used to specify doses of medicines and refer to date and time. The associations reflect new methods in eighteenth-century medicine: collocates on inoculating the smallpox are important, and the effects of inoculation are evaluated by statistical information. Thus, the eighteenth-century developments in medicine and scientific thinking affect the representation of patients. They are seen more often as objects of scientific inquiry and related to similar cases with numerical data, although the everyday practices of treatment remain at the centre stage.

The patients are often presented as objects of medical treatments, but there are also many instances of patients portrayed as active participants. The active role is more notable in works aimed at lay readers than in the professional texts, since they give practical health advice on a regimen and recipes and also advise self-inoculating the smallpox, for instance. The corpus text *Duty of Physicians and Patients* (1715: 33) by Parker on medical ethics instructs patients to actively participate in the consultation with physicians by carefully narrating their symptoms (e.g. *When the Physician's come, 'tis the Patient's Duty* […] *to represent his Condition as distinctly and fully to him as he can*). In general, the patient's role as a more active participant became more notable in the era, as patients started to share their own experiences in *The Gentleman's Magazine* (see Taavitsainen 2019, Porter 1992), and there are further testimonies written by patients on the effects of cures within texts in other categories of LMEMT. In contrast, Fissell (1991: 148–152) notes that the patients' narratives of their own illnesses became less important in infirmaries: the practitioners increasingly relied on physical observation of patients, as their education improved along with professionalisation of medicine. Porter (1992: 91) summarises that the patient's role was complex in medical consultations in the era since "the sick have not just been 'patients' but 'agents' as well, both looking after their own health, and playing active roles in managing their dealings with medical professionals and the institutions of regular medicine." The patients actively participated in their cure, which emerges particularly in the texts for lay audiences.

The representation of patients differs according to target audience, the topic of texts and corpus category, and the medical authors view patients with sympathy especially in many texts for lay audiences. Compassion can be seen as an important component of eighteenth-century lay medical writing surrounding patients, and it was likely affected by the polite style and philanthropic spirit of the era. This compassionate view and taking care of the sick and poor was useful for the upper layers of society: it protected workers from diseases and prevented contagious diseases from spreading to the upper classes (Porter 1995: 33). The compassionate attitude

expressed towards lay readers benefited the physicians as well. By taking part in public correspondence, the physicians were able to demonstrate their medical skills to affluent readers and promote their career (Porter 1985: 313). The voluntary hospitals provided public arenas for polite interaction: the institutions aided the poor but at the same time the aim was to entertain more affluent visitors to attract subscriptions (Klein 2002: 895). The hospitals were places for the physicians to mingle with the upper classes, and the importance of their reputation surfaces especially in the descriptive collocates such as *eminent physician*. The physicians aspired to improve the status of their profession by emphasising the role of the College of Physicians and by attacking apothecaries, which is visible in many texts in LMEMT. The sympathetic view towards patients surfaces in the polemical texts on physicians versus apothecaries in addition to the lay texts. The motivation in these contexts is, nevertheless, different, as the purpose is to attack the apothecaries by portraying the patients as victims of their treatments.

The medical practitioners are typically linked to actions, other practitioners and quantities, and the collocates describe the attributes of professionals. The collocates hence mainly focus on the practitioners and practicing of medicine, and a qualitative assessment reveals that these topics remain in LMEMT across the corpus timeline. The authors refer to the writings of others and use mental verbs, and the concept of knowledge is repeated. The physicians are regularly related to the College of Physicians and the healing art, emphasising the role of education in their profession. A recurrent theme is to describe (especially educated) physicians and surgeons by positive attributes, and many of these wordings emerge as formulaic sequences. About 13.1% of collocate types for physicians and 9.7% for surgeons are overt descriptive references to the practitioners, and they generally highlight positive characteristics, and similar positive collocates are found among the other semantic categories. Physicians can, therefore, be related to the semantic prosody of praise that characterises their shared professional identity, i.e. they are associated positively with proper skills, reputation and education. The positive descriptions surround surgeons as well, which can be seen as somewhat unexpected, as the surgeons ranked lower in their professional status compared to the physicians. The surgical texts, nevertheless, have a long history in medical writing, and surgery was often the leading field of new methods in medicine, improving the representation of surgeons.

For apothecaries, the same semantic categories generally prevail as for physicians and surgeons, the focus being on the practitioners, i.e. actions, references to practitioners, quantities and abstract concepts. However, most of the collocates indicate critical descriptions. Therefore, the characterisation of apothecaries is overall noticeably more negative, and the lack of collocates expressing positive descriptive

references poses an especially distinct contrast between physicians/surgeons and apothecaries. The apothecaries are overtly described as *ignorant* and they are said to *pretend* when practicing medicine. Across the topics, apothecaries are blamed for their fees and businesses in an unfavourable light (e.g. *wages, business, gains, trade, bill* and *salary*). In fact, the apothecaries are associated with the desire to solely make profit instead of being interested in the wellbeing of patients. References to money are more expected considering that the poor visited apothecaries for treatment because of their lower fees.

The apothecaries are not portrayed treating patients successfully or even preparing medicines in LMEMT. Loudon (1992: 225) states that the standing of the lower status medical practitioners improved at the end of the eighteenth century. This type of change in status is not detected in the representation surrounding apothecaries, as the semantic prosody remains negative. The critical portrayal of apothecaries is visible especially in the texts that address the controversy between physicians and apothecaries, but prevails in other contexts as well. However, this negative representation may not be fully representative of the extratextual reality. An increasing number of people were able to use medical services, encouraged by the optimistic attitude of the Enlightenment and an improved financial status of patients (Porter 1995: 25, 30). Nevertheless, the patients did not hesitate to seek consultation from several types of practitioners or use quack medicines, even though the apothecaries were openly criticised in the eighteenth century (see Corfield 2009: 4). Against this sociohistorical background, it can be seen that the medical genre likely presents the apothecaries in a more negative light than many other genres. This reflects the idea that some meanings emerge only in specific genres (see e.g. Hunston 2007).

The representation of the physicians can also be compared to the patients' writings in eighteenth-century letters and diaries (see Lane 1985: 218–233). In these texts, the patients rarely criticise their medical practitioners apart from occasional comments on their wrong diagnoses and fees. On the other hand, the patients record the preference of some physicians to treat more affluent patients in the hope that it enhanced their reputation. These associations seem to be more in line with the positive idea of physicians in LMEMT.

Overall, the representations of patients and practitioners are very different in the medical texts, as the discussion on patients generally concentrates on their treatment, while the descriptions of practitioners focus on their skills and on reputation. In his *Directions Tending to Health* (1717: 8), Byfield succeeds in summarizing the representations of practitioners and patients: he notes that when patients recover, *the Dr. will have the Credit, and the Patient the Benefit.*

References

Anthony, Laurence. 2014. AntConc. Computer software. Tokyo, Japan: Waseda University. www.laurenceanthony.net

Baker, Paul. 2014. *Using Corpora to Analyze Gender*. London: Bloomsbury.

Baker, Paul, and Tony McEnery. 2005. "A Corpus-Based Approach to Discourses of Refugees and Asylum Seekers in UN and Newspaper Texts." *Journal of Language and Politics* 4 (2): 197–226. https://doi.org/10.1075/jlp.4.2.04bak

Bäcklund, Ingegerd. 2006. "Modifiers Describing Men and Women in Nineteenth-Century English." In *Nineteenth-Century English: Stability and Change*, ed. by Merja Kytö, Mats Ryden, and Erik Smitterberg, 17–55. Cambridge: Cambridge University Press.

Bell, Allan. 2001. "Back in Style: Reworking Audience Design." In *Style and Sociolinguistic Variation*, ed. by Penelope Eckert, and John R. Richford, 139–169. Cambridge: Cambridge University Press.

Brezina, Vaclav, Tony McEnery, and Stephen Wattam. 2015. "Collocations in Context: A New Perspective on Collocation Networks." *International Journal of Corpus Linguistics* 20 (2): 139–173. https://doi.org/10.1075/ijcl.20.2.01bre

Burr, Vivien. 2002. *The Person in Social Psychology*. New York: Psychology Press.

Corfield, Penelope J. 2009. "From Poison Peddlers to Civic Worthies: The Reputation of the Apothecaries in Georgian England." *Social History of Medicine* 22 (1): 1–21. https://doi.org/10.1093/shm/hkn096

Fissell, Mary E. 1991. *Patients, Power and the Poor in Eighteenth-Century Bristol*. Cambridge: Cambridge University Press.

Gabrielatos, Costas, and Paul Baker. 2008. "Fleeing, Sneaking, Flooding: A Corpus Analysis of Discursive Constructions of Refugees and Asylum Seekers in the UK Press, 1996–2005." *Journal of English Linguistics* 36 (1): 5–38. https://doi.org/10.1177/0075424207311247

Gries, Stefan Th. 2013. "50-Something Years of Work on Collocations: What Is or Should Be Next…" *International Journal of Corpus Linguistics* 18 (1): 137–166. https://doi.org/10.1075/ijcl.18.1.09gri

Gupta, Kat. 2015. *The Representation of the British Suffrage Movement*. London: Bloomsbury.

Hiltunen, Turo, and Jukka Tyrkkö. 2009. "'*This well known to barbers and laundresses*: Overt References to Knowledge in English Medical Writing from the Middle Ages to the Present Day." In *Corpus Linguistics: Refinements and Reassessments*, ed. by Antoinette Renouf, and Andrew Kehoe, 67–86. Amsterdam: Rodopi.

Hunston, Susan. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139524773

Hunston, Susan. 2007. "Semantic Prosody Revisited." *International Journal of Corpus Linguistics* 12 (2): 249–268. https://doi.org/10.1075/ijcl.12.2.09hun

Klein, Lawrence. 2002. "Politeness and the Interpretation of the British Eighteenth Century." *The Historical Journal* 45 (4): 869–898. https://doi.org/10.1017/S0018246X02002716

Lane, Joan. 1985. "'The doctor scolds me': The Diaries and Correspondence of Patients in Eighteenth Century England." In *Patients and Practitioners: Lay Perceptions of Medicine in Pre-Industrial Society*, ed. by Roy Porter, 205–248. Cambridge: Cambridge University Press.

Lehto, Anu. 2019a. "Changing Portrayals of Medicine and Patients in Eighteenth-Century Medical Writing: Lexical Bundles in PUBLIC HEALTH, METHODS and Case Studies." In *Late Modern English Medical Texts. Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen, 145–172. Amsterdam: John Benjamins. https://doi.org/10.1075/z.221.08leh

Lehto, Anu. 2019b. "The Representation of Citizens and Monarchy in Acts of Parliament in 1800 to 2000: Identifying Social Roles through Collocations." In *Corpus-Based Research on Variation in English Legal Discourse*, ed. by Teresa Fanego, and Paula Rodríguez-Puente, 235–260. Amsterdam: John Benjamins. https://doi.org/10.1075/scl.91.10leh

Loudon, Irvine. 1992. "Medical Practitioners 1750–1850 and the Period of Medical Reform in Britain." In *Medicine in Society: Historical Essays*, ed. by Andrew Wear, 219–248. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511599682.008

Louw, Bill. 1993. "Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies." In *Text and Technology: In Honour of John Sinclair*, ed. by Mona Baker, Gill Francis, and Elena Tognini-Bonelli, 157–176. Amsterdam: John Benjamins. https://doi.org/10.1075/z.64.11lou

McEnery, Tony, and Helen Baker. 2019. "Language Surrounding Poverty in Early Modern England: A Corpus-based Investigation of How People Living in the Seventeenth-century Perceived the Criminalized Poor." In *From Data to Evidence in English Language Research*, ed. by Carla Suhr, Terttu Nevalainen, and Irma Taavitsainen, 223–257. Leiden: Brill.

McEnery, Tony, and Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

Partington, Alan. 2004. "'Utterly content in each other's company': Semantic Prosody and Semantic Preference." *International Journal of Corpus Linguistics* 9 (1): 131–156. https://doi.org/10.1075/ijcl.9.1.07par

Pearce, Michael. 2008. "Investigating the Collocational Behaviour of *man* and *woman* in the BNC Using Sketch Engine." *Corpora* 3 (1): 1–29. https://doi.org/10.3366/E174950320800004X

Porter, Roy. 1985. "Laymen, Doctors and Medical Knowledge in the Eighteenth Century: The Evidence of *The Gentleman's Magazine*." In *Patients and Practitioners: Lay Perceptions of Medicine in Pre-Industrial Society*, ed. by Roy Porter, 283–314. Cambridge: Cambridge University Press.

Porter, Roy. 1992. "The Patient in England, c. 1660–c. 1800." In *Medicine in Society: Historical Essays*, ed. by Andrew Wear, 91–118. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511599682.004

Porter, Roy. 1995. *Disease, Medicine and Society in England, 1550–1860*. 2nd edition. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511608117

Ratia, Maura. 2020. "Finding Evidence for Changing Society: A Collocational Study of Medical Discourse in 1500–1800." In *Corpora and the Changing Society: Studies in the Evolution of English*, ed. by Paula Rautionaho, Arja Nurmi, and Juhani Klemola, 57–78. Amsterdam: John Benjamins. https://doi.org/10.1075/scl.96.03rat

Reill, Peter Hanns. 2003. "The Legacy of the 'Scientific Revolution': Science and the Enlightenment." In *The Cambridge History of Science: Volume 4, Eighteenth-Century Science*, ed. by Roy Porter, 23–43. Cambridge: Cambridge University Press. https://doi.org/10.1017/CHOL9780521572439.003

Risse, Guenter B. 1992. "Medicine in the Age of Enlightenment." In *Medicine in Society: Historical Essays*, ed. by Andrew Wear, 149–196. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511599682.006

Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, John. 2004. *Trust the Text: Language, Corpus and Discourse*. London: Routledge. https://doi.org/10.4324/9780203594070

Sinclair, John, Susan Jones, and Robert Daley. 2004. *English Collocation Studies: The OSTI Report*, ed. by Ramesh Krishnamurthy. London: Continuum.

Stubbs, Michael. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.

Stubbs, Michael. 2003. "Two Quantitative Methods of Studying Phraseology in English." *International Journal of Corpus Linguistics* 7 (2): 215–244. https://doi.org/10.1075/ijcl.7.2.04stu

Sveen, Hanna Andersdotter. 2010. "Good-Natured Fellows and Poor Mothers: Defining Social Roles in British Nineteenth-Century Children's Literature." In *Social Roles and Language Practices in Late Modern English*, ed. by Päivi Pahta, Minna Nevala, Arja Nurmi, and Minna Palander-Collin, 211–228. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.195.10sve

Taavitsainen, Irma. 2019. "Polite Society Language Practices: Letters to the Editor in *The Gentleman's Magazine*." In *Late Modern English Medical Texts. Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen, 129–144. Amsterdam: John Benjamins. https://doi.org/10.1075/z.221.07taa

Taavitsainen, Irma, and Turo Hiltunen (eds). 2019. *Late Modern English Medical Texts. Writing Medicine in the Eighteenth Century*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.221

Taavitsainen, Irma, Turo Hiltunen, Anu Lehto, Ville Marttila, Päivi Pahta, Maura Ratia, Carla Suhr and Jukka Tyrkkö. 2019. *Late Modern English Medical Texts: The Corpus*. In *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.221

Taavitsainen, Irma, Peter Murray Jones, and Turo Hiltunen. 2019. "Sociohistorical and Cultural Context of *Late Modern English Medical Texts*." In *Late Modern English Medical Texts. Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen, 17–30. Amsterdam: John Benjamins. https://doi.org/10.1075/z.221.02taa

Taavitsainen, Irma, Päivi Pahta, and Martti Mäkinen (eds). 2005. *Middle English Medical Texts*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.131

Taavitsainen, Irma, Päivi Pahta, Turo Hiltunen, Martti Mäkinen, Ville Marttila, Maura Ratia, Carla Suhr, and Jukka Tyrkkö. 2010. *Early Modern English Medical Texts: Corpus*. In *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen, and Päivi Pahta. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.160

Withey, Alun. 2019. "Household Medicine and Recipe Culture in Eighteenth-Century Britain." In *Late Modern English Medical Texts. Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen, and Turo Hiltunen, 113–128. Amsterdam: John Benjamins. https://doi.org/10.1075/z.221.06wit

Wray, Alison. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519772

Xiao, Richard. 2015. "Collocation." In *The Cambridge Handbook of English Corpus Linguistics*, ed. by Douglas Biber, and Randi Reppen, 106–124. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139764377.007

The discursive dynamics of personal experience narratives and medical advice in 18th-century British consultation letters

The case of Dr. William Cullen

Anna Franca Plastina

University of Calabria

Medical consulting by letter was a common practice in eighteenth-century Britain since physical examination was still far from being an established diagnostic tool. Despite what this epistolary practice testifies of the importance of illness narratives in historical medical discourse, scant attention has been devoted to their discursive representations. This chapter examines the discursive dynamics of medical consulting letters, drawing on the Labovian framework and the RIAS (Roter Interaction Analysis System) method for coding medical interaction. Discourse analysis is performed on a dataset of sample letters from Dr. Cullen's (1710–1790) collection to unpack the referential and evaluative functions of narratives by patients, and Cullen's discourse strategies for building doctor-patient relationship. Results from the case study highlight the historical significance of mediating and interpreting illness narratives according to the Enlightenment symptom-based model.

Keywords: illness narratives, discursive dynamics, personal experience narratives (PENs), consultation letters, referential and evaluative functions, 18th-century medical epistolary discourse, William Cullen

1. Introduction

Early modern humoral medicine appeared to exert a strong influence on the way in which the procedure of diagnosis was perceived and performed by eighteenth-century British physicians. Very limited relevance was, in fact, accorded to the patient's physical examination in the tradition of humoral medical theory, whereby diseases were attributed to an imbalance of the four humors existing as body liquids, namely blood, phlegm, black bile and yellow bile. Learned

physicians further relied on this theoretical approach to claim their intellectual role as "thinkers" compared to the manual work of less-learned surgeons considered as "touchers" (Porter and Porter 1989: 75). Thus, the hands-on physical examination was far from an established diagnostic tool (Wild 2000); patient accounts were therefore crucial for diagnostic decisions in Enlightenment medical practice since only patients as the initial sufferers had direct access to their own symptoms. Accordingly, "the physician […] arrived at a more precise diagnosis thanks to his skill in interpreting the patient narrative" (Stolberg 2013: 100). Even if the patient's body was occasionally observed, Enlightenment medical practice was largely based on "verbal testimony, such that the practitioner did not always have to see his patients, relying instead on a written report to formulate his opinion: hence the relative importance of doctor-patient correspondence during the eighteenth century" (Louis-Courvoisier and Pilloud 2004: 71).

Medical consulting by letter was a common practice in eighteenth-century Britain not only owing to the influence of the Galenic tradition on the medical approach to diagnosis, but also due to other sociohistorical factors. While letter-writing was a constituent part of physicians' professional and learned activities (Steinke and Stuber 2004: 142), the expansion of the eighteenth-century British postal service further facilitated this practice (Overton 2007). In addition, literate sufferers frequently resorted to consulting by letter either due to the lack of physicians in their area or to obtain expert advice from famous doctors, or even to directly seek second opinions (Louis-Courvoisier and Mauron 2002). Moreover, the British were "well-versed in medical matters" as patients kept abreast of both "the current state of medical knowledge" and of "the fashionable jargon of their physicians" (Wild 2006: 19). The common eighteenth-century reader, in fact, could easily find out about "remedies in general interest publications such as almanacs as well as in volumes specifically devoted to medical advice […]" (Porter 1985a: 140). Hence, patients and doctors shared "a common rhetoric that facilitated medicine-by-post", which made a significant contribution to curtailing professional authority exercised through "the power of an exclusive language" (Wild 2006: 19).

As an effective and convenient mode of communication, medical consulting by letter testifies to the importance of illness narratives in eighteenth-century medical discourse. Illness narratives belong to the genre of personal experience narratives (PENs), which relate an event from the first-person perspective marked by a sense of ownership compared to other popular narrative genres. In general, illness narratives are considered "autobiographical accounts of illness spoken or written by patients" (Jurecic 2012: 2). The eighteenth-century written account can be particularly considered a "pathography", or "an extended narrative situating the illness experience within the author's life and the meaning of that life […] [It] tends to focus on the emotional components of a medical experience […] [and] to dramatize the

events of illness" (Hawkins 1999: 13). Hence, the practice of "eighteenth-century medicine by post" reveals a new rhetoric of sensibility fostering "patients' narration of their case histories as *experiential*, not merely as a compendium of exact physical symptoms" (Wild 2006: 10; emphasis added).

In spite of its historical significance in eighteenth-century British medical discourse, this peculiar kind of correspondence has been mostly considered from the medical perspective to reflect on disease treatment (Steinke and Stuber 2004); scant attention has been devoted to the function of PENs in epistolary representations of historical medical discourse.

This chapter attempts to fill this void by exploring the discursive dynamics of PENs and medical responses in the eighteenth-century epistolary exchanges of Dr. William Cullen (1710–1790; Royal College of Physicians of Edinburgh), whose epistolary opinion was in high demand at the time. The case study looks at how British sufferers mediate the key referential and evaluative functions of their PENs, and also how Cullen strategically constructs doctor-patient relationships in his responses. From a socio-discursive perspective, a detailed analysis of these representations may offer deeper insight into this epistolary practice as a representative source of eighteenth-century British medical discourse.

2.   Orientation and evaluation in personal experience narratives

Orientation and evaluation are considered to be two key components in shaping meaning in PENs (Labov 1972). Orientation provides information about the "place, time, and behavioral situation" (Labov and Waletzky 1997: 32) to ground the narrative in the contextual world by means of its *referential* function in discourse. Thus, "the illness is articulated and positioned in time and space, and within the framework of a personal biography" (Hydén 1997: 56). As a referential analytic device, time aids the understanding of the temporal-historical perspective of illness; space helps shed light on how the body as the centre of experience is considered from a phenomenological perspective. Time and space deictics can be "absolute", i.e., "to locate events in 'absolute' time relative to some absolute *origo*", or "relative", i.e., "the time at which the events in the narrative occurred" and "the place of utterance" (Levinson 1983: 67, 73). *Absolute space* is represented through spatial discourse markers (e.g. spatial deixis, place names) referring to the sociohistorical context framing the narrative; *absolute time* through temporal discourse markers (e.g. temporal deixis, time expressions) pointing to time in the outer world; *relative space* here indicates specific body parts which locate the sufferer's narration of physiological distress, while *relative time* is circumscribed to temporal expressions regarding the personal illness itself. Relative referentials have an important function

as they offer insight into the dynamic structure of PENs. Relative time particularly allows to classify PENs as *progressive, stability* or *regressive* narratives (Gergen and Gergen 1986), and thereby to establish the illness trajectory.

Unlike the contextualizing function of orientation, evaluation gives the text its significance and establishes the narrative point (Labov and Waletzky 1997). Hence, "evaluative clauses (and evaluative aspects of referential clauses) have to do with why the narrator is telling the story" (De Fina and Johnstone 2015: 153). Labov (1972: 373–374) identifies four main functional types of evaluation: *external evaluation*, when narrators provide their interlocutor with the reason(s) for telling the story, and thus recount events from a position outside the story; *embedded evaluation* ensures the dramatic continuity of the narrative from within since the tellers describe their feelings at the time of the event; *evaluative action* is used by the tellers as a means of "dramatizing the evaluation of a narrative" through the actions taken; and *evaluation by suspension* involves "stopping the action" to draw attention to a specific part of the narrative. Accordingly, four main kinds of syntactic devices may be used to express different evaluative types (Labov 1972: 378–390):

1. *intensifiers* (e.g. quantifiers, repetition, deictics) to strengthen narrative events;
2. *comparators* (e.g. modals, negatives, quasimodals, superlatives, comparatives) as more complex devices to "compare the events which did occur to those which did not occur" (Labov 1972: 381);
3. *correlatives* (e.g. progressives such as *be + V-ing*, appended participles as sequence of verbs with *-ing*, double attributives), which "bring together two events so that they are conjoined";
4. *explicatives*, or appended subordinate clauses introduced by conjunctions (e.g. *while, although, since*), which justify the main narrative.

In the specific case of illness narratives, referential and evaluative functions are meant to aid the physician's interpretation of PENs, thus allowing for a more precise diagnosis. While the evident goal of the epistolary response is to provide tailored medical advice, it can be envisaged that doctor-patient relationship is also mediated in this kind of socio-medical communication. A useful method to capture the different discourse strategies serving this twofold purpose in epistolary medical responses is provided by the Roter Interaction Analysis System (RIAS) (Roter and Larson 2002). Although the RIAS is used worldwide to code current medical exchanges, its application to historical medical exchanges has not been taken into account so far. Yet, the RIAS categories appear to be equally suitable for coding discourse strategies in historical medical responses. The RIAS classifies discourse strategies as:

i.   *task-focused strategies* for medical purposes, which include the two subcategories of *data-gathering* (biomedical questions) and *patient education* (biomedical topics);

ii. *socio-emotional strategies* for building doctor-patient relationship, which comprise the four subcategories of *social, positive, negative* and *emotional talk*;

iii. *socio-emotional strategies* for *facilitation* (e.g. asking for patient opinion) and *patient activation*, which includes *procedural talk* (e.g. providing medical instructions).

3.   The case of Dr. William Cullen's epistolary collection

3.1   Aims and method

The case study addresses the following two research questions:

RQ1:   How do eighteenth-century British sufferers mediate the referential and evaluative functions of their PENs in consultation letters addressed to Dr. Cullen?

RQ2:   Which discourse strategies does Dr. Cullen employ in response to these PENs to build doctor-patient relationship?

Following Fairclough (1992: 230), a mixed-method design is used to perform a "detailed analysis of a small number of discourse samples", which allows to keep the research within manageable bounds and thus facilitates in-depth analysis.

3.2   Materials

Cullen's private consultation letters survive as a remarkable archive of his postal medical correspondence from the mid-1750s to 1790. The numerous letters include requests for advice from patients coming from different social classes and geographical locations along with copies of Cullen's responses. The digitized collection (www.cullenproject.ac.uk/) represents a rare case of the survival of both sides of eighteenth-century epistolary exchanges, and thus provides a rich resource for gaining historical insight into the experience of illness and the effort to cure disease. A total of 23 sample letters documenting four distinct epistolary cases were downloaded from the digitized collection. The selection criteria included letters featuring first-person PENs of different illnesses with a cut-off of two epistolary exchanges (four letters) to ensure unfolding narratives. The collected materials (17,997 words)

include 11 PENs (8,767 words) and 12 responses (9,230 words) by Dr. Cullen with the epistolary exchanges spanning over more than a decade (1774–1787). The labels attributed to each of the four cases are the work of the Cullen project. In detail, Case 1 (www.cullenproject.ac.uk/case/56/) consists of 5 letters (2,617 words) written between the MP Anthony Chamier and Cullen in 1780. The patient suffers from general "languor" or weakness and severe stomach pain, which unfortunately prove fatal. The digital collection labels this case as "A Journey in Search of Health". After consulting Cullen in Edinburgh, Chamier travels south to London, via Buxton and Matlock (Derbyshire), keeping the physician informed of his failing condition along the way. Cullen had planned the route to enable Chamier to be in the vicinity of Tamworth (Staffordshire) at the time of what proved his re-election to Parliament. Case 2 (www.cullenproject.ac.uk/case/12/) includes 4 letters (1,446 words) exchanged between Mrs. Jane Webster from York and Cullen between 1780 and 1781. The patient seeks advice on losing weight to improve her health as a case of "The Modern Concern About Obesity". Case 3 (www.cullenproject.ac.uk/case/369/) refers to Mr. Robert Ligertwood from Aberdeen, who believes to have a "nervous weakness" due to "A Life of Intemperance". 7 letters (4,527 words) document this doctor-patient correspondence which spans between 1777 and 1781. Finally, Case 4 (www.cullenproject.ac.uk/case/443/) involves Mr. Henry Lochhead, who regularly travels between Virginia (USA) and Glasgow, and has a genito-urinary complaint which seems to be venereal. "The Business Traveller Abroad" and Cullen exchange 7 letters (9,407 words) between 1774 and 1787. The specialized collection is thus heterogeneous in its representation of the patients' different social status, gender, place of origin and pathologies. Additionally, its small size appears adequate to address the research questions posed, and to perform a detailed manual annotation of the linguistic features under consideration, which seem to be frequently instantiated in epistolary medical consultations (cf. McEnery et al. 2006: 72).

3.3   Procedure

The images of the letters were downloaded as raw data from the aforementioned digitized collection, manually transcribed as a workable sample collection for present research purposes and analyzed in three distinct phases. In the first, referentials were treated as the Labovian orientation component based on the parameters of spatiotemporality (time, space, time-space) and its types (absolute, relative). Linguistic expressions were manually annotated and coded as absolute time (e.g. *in 1775*), relative time (e.g. *I had began two days before*), absolute space (e.g. *at Matlock*), relative space (e.g. *in my Stomach*), combined absolute/relative time-space (e.g. *In November 1777, I went by sea to London, and aboard I was sized in the Morning with such dreadful simptoms from a Weight in my head*).

The dynamic structure of the PENs was then identified by applying Gergen and Gergen's (1986) classification to establish the illness trajectory based on the type of narrative, namely progressive, stable or regressive.

In the second phase, evaluative discourse markers were manually coded, based on the most suitable Labovian evaluation types (external, embedded, evaluative action) and syntactic devices (intensifiers, comparators, correlatives). External evaluation was considered as patients' direct interaction with Dr. Cullen outside the illness narrative (*I am so much obliged to you Sir for your very Clear and Satisfactory Opinion Concerning my Case*); embedded evaluation as the dramatic continuity of the narrative from within (*I suffered exceedingly from the spasms, and of course am far more cautious*); and evaluative action as the remedies taken (*I followed rather a more rigid Plan, I diminishd the Quantity of food*).

In the third and final phase, the two RIAS categories (task-focused and socio-emotional) were applied to Cullen's epistolary responses to code his discourse strategies. Two task-focused subcategories were used to code Cullen's medical questions (data-gathering) and information-giving (patient education); four socio-emotional subcategories were applied to linguistic expressions of nonmedical responses (social talk), approval or agreement (positive talk), disagreement or criticisms (negative talk), concerns, reassurance and empathy (emotional talk) as representations of doctor-patient relationships. While different strategies may occur within the same phrase, these were treated separately for present research purposes. Two other socio-emotional subcategories were further introduced to code Cullen's requests for patient opinion (patient activation) and his provision of medical instructions (procedural talk).

4. Results and discussion

4.1 The referential functions of the PENs

A marked combination of space-time referentials was found to characterize the orientation component of all the PENs in the collection:

(1) It is now <u>ten days since</u> we left <u>Edinburgh</u>; I had began <u>two days before</u> to diminish the absorbent[1] and some fermentation had arisen probably in consequence of it, <u>in my Stomach</u>.   (Anthony Chamier, 1780)

---

1. Medicine employed to absorb acid, usually "certain earths suited to take acids into their pores, and at the same time to destroy their acid quality" (www.cullenproject.ac.uk/data/items/actions/).

As underlined, the non-calendrical measure of time (ten days since) and the spatial marker of place name (Edinburgh) construct the social context and thus function as absolute referentials for the fixed point of interest, namely the sufferer's bodily spatiotemporality of illness, marked by relative referentials (two days before; in my stomach). Absolute referentiality is thus meant to allow Dr. Cullen (and the reader in general) to construct initial mental representations of the social context within which the narrated health complaint is developed.

Once the sociohistorical frame is constructed by means of absolute spatiotemporal referentials, it is replaced by the more important illness frame of the "text world" of the narrative mediated mainly by relative spatiotemporality. This explains why the occurrence of relative spatiotemporal referentials ($N = 60$) is almost double that of absolute ones ($N = 31$). This dynamic "priming" process of "frame switch" (Emmott 1997: 147) thus operates to focus the reader's attention on the illness narrative as the core component of PENs in consultation letters:

(2)   a.   I returned to <u>London</u> and from <u>thence</u> by sea to <u>Aberdeen</u> where I staid for <u>six months.</u>
      b.   For <u>2 or 3 years past</u> I could not wear a cockt hatt otherwise my <u>eyes and brain</u> would be so affected with a Giddiness.   (Robert Ligertwood, 1778)

The frame switch between absolute spatiotemporality referring to the social context in Example (2a.) and relative spatiotemporality referring to the "illness world" in Example (2b.) shows how the referentials create a dynamic change, whereby the illness frame is "primed" so that Giddiness receives the main focus of attention.

Despite the fact that a consistently lower frequency of absolute referentials ($N = 72$) was recorded for all the three referential categories (time, space, time-space) compared to their relative counterparts ($N = 118$), their systematic use appears to be beyond the simple function of setting up the initial social frames of PENs. This suggests that absolute referentials in the background frames tend to act also as a "bound" throughout the narrative (Emmott 1997: 123). In other words, absolute referentials help construct the "contextual frame" within which the sufferers are constantly "bound" in specific sociohistorical contexts throughout their PENs.

This functionality is further understood to operate as a form of "contextual monitoring" (Emmott 1997), which enabled Dr. Cullen as the direct reader to constantly track any possible health changes in the sufferer. Contextual monitoring proves to be particularly helpful, as the epistolary practice, by its very nature, does not enable immediate verbal feedback. Ultimately, this referential strategy contributes to monitoring the evolving health condition, thus allowing Dr. Cullen to better assess the course of the illness as progressive, stable or regressive (Gergen and Gergen 1986). This is skillfully accomplished by all four narrators themselves as they discursively engage in the contextual monitoring of their own illness trajectories:

(3) a. I see with terror the season advancing and that <u>I make no progress</u>; Before I rise it seems that I have alacrity and vivacity (Anthony Chamier, 15th Sept. 1780); but I have no sooner been put to the task of bearing my own weight, <u>then all is gloom, languor, and sloth</u>.   (27th Sept. 1780)

  b. As soon as the warm weather came in I was advised into the country, <u>my Cough intirely left me</u>, and in a short time <u>I was quite free from it</u> (Jane Webster, 1780); I found that in two months <u>I had lost several pounds</u>.
  (1781)

  c. <u>my breathing seems gratly better</u> for I cant exceed above a bottle of claret after dinner (Robert Ligertwood, 1777); Upon my arrival at London, where I staid abowt six weeks I felt <u>a return of the same simptoms and distress both in body and mind</u> (1778); A few days since he has drawn from my Arm about eleven ounces blood by which I found <u>a little present benefite, but am now much as formerly</u> (1780); since writing […] <u>I have found my Stomach so uneasy</u>.   (1781)

  d. <u>some other symptoms</u> which happened since this morning and which I do not remember ever to have experienced before have determined me [to send this letter] (Henry Lochhead, 1777); as in all this time which was about 6 weeks <u>I was in the best health</u> I hads been in a long time (Jan. 1778); <u>I found the most sensible good effects</u> in the urethra by month of April I found <u>I had gained considerably</u>.   (Feb. 1778)

As underlined in all instances, the narrators explicitly track their own health changes by using a combination of absolute and relative referentials which provide Dr. Cullen with meaningful contextual clues to aid his reading of the letters and facilitate his advice-giving responses. In detail, the instances from the two letters in Example (3a.) show how the initial "stability" narrative (*I make no progress*) is afterwards transformed into a regression narrative (*all is gloom, languor, and sloth*). In Example (3b.), instead, the narratives are equally progressive in both letters (*I was quite free from it; I had lost several pounds*). On the other hand, Example (3c.) shows how the narrator's condition is monitored as constantly changing across the four letters, oscillating between progression (*my breathing seems gratly better*), regression (*a return of the same simptoms and distress*), progression (*a little present benefite*), stability (*now much as formerly*) and regression again (*my Stomach so uneasy*). Example (3d.) indicates how the initial regression narrative (*some other symptoms*) in the first letter is later transformed into progression narratives in the following two (*I was in the best health; the most sensible good effects; I had gained considerably*).

Interestingly, the contextual clues provided by the various referential devices offer an insight into both the sociohistorical context and the medical practices at the time (see also Examples (5) and (6)):

(4) a.   The situation of my affairs makes it necessary for me to go to <u>America</u>, notwithstanding the disorder that yet prevails in <u>that country</u>.   (1778)

    b.   I am told of there being no other care beside the <u>Bougie</u> […] I might find too great difficulty in getting one past. (this last reason woud be stranger in its effect on me if I were in <u>Virginia</u> where I caud not get so good assistance)   (Henry Lochhead, 1777)

The spatial markers (*America*, *that country*) in Example (4a.) implicitly refer to the disorder caused by the American War of Independence (1775–1783), whereby the Thirteen Colonies won independence from Great Britain, thus becoming the United States of America. These historical clues are offered by Henry Lochhead, who had to travel to Virginia on business and was experiencing urinary problems due to the use of a bougie. Thus the reader is offered insight into the sufferer's understanding that the practice of bougies, or cloth "tubes" used for surgical drainage, was the only treatment available at the time for genito-urinary problems.[2] This is, in fact, confirmed in Cullen's clear response: "when the disease is more recent the most certain cure is by Bougies but when of longer standing and especially when it has effected the neck of the bladder the Bougies are not so easily admitted and some times not so successfully".[3] In Example (4b.), Lochhead further informs the reader that medical assistance for this treatment appeared not to be as effective in Virginia as it was in Great Britain.

The absolute temporal referential (during Winter) in Example (5a.) helps create the contextual frame within which the narrator reports about the practice of induced vomiting (*I used my finger or a feather*) for the purpose of cleansing the stomach of phlegm. Likewise, the temporal referential (*a few days since*) in Example (3c.) sets the context for representing the practice of bloodletting to expel the corrupt humors from the body.[4] In Example (5b.), the temporal marker of duration refers, instead, to the lengthy practice of purging with mana (a fruit sugar usually derived from the sugary sap of the tree Fraxinus ornus) for the purpose of cleansing the body of impurities, usually by stool.[5]

(5) a.   Doctor Livingston frequently calls upon me and I have by his advice, <u>during Winter</u> taken three or four <u>Vomits</u> which never wrought well till I used my finger or a feather.   (Robert Ligertwood, 1780)

---

2.   www.cullenproject.ac.uk/data/items/treatments/

3.   www.cullenproject.ac.uk/docs/3694/

4.   The drawn blood was usually left in a dish to cool before being examined by a surgeon or physician as a form of a diagnostic test.

5.   www.cullenproject.ac.uk/items/i76/

b. Well, I went on in this manner (but purged with manna) <u>for 20 days</u> – and
    I may thank God you directed the manner of purge proper.

<div align="right">(Henry Lochhead, 1787)</div>

Referentials also point explicitly to pre-nineteenth century humoral medical theory
(6):

(6) but on riding <u>in the mornings</u> I find the Giddiness <u>in my Head</u>, which I attribute
     to <u>the Humours</u> being sett in motion.   (Robert Ligertwood, 1780)

The sufferer appears to attribute the condition of giddiness to an imbalance of the
four bodily fluids (humours) through the use of the relative spatial referential (in my
head). The absolute temporal referential (in the mornings) informs the reader of the
narrator's daily habit of horse-riding, a practice which was clearly recommended by
Cullen in his clinical lectures as a method of treatment for hypochondria.[6]

On the whole, the significant occurrence of spatiotemporal referentials ($N = 190$
tokens) (see Table 1) generally highlights the great importance of referentiality. In
shaping PENs as factual narratives, which advance "claims of referential truthful-
ness" (Schaeffer 2014: 179; original emphasis), the vivid referential details of facts
help reconstruct the primacy of the personal illness experience, thereby allowing
the sufferer to gain credibility.

Table 1. Distribution of the Referential Functions in the sample PENs

Referential function ($N = 190$)	Case 1 ($N = 44$)	Case 2 ($N = 28$)	Case 3 ($N = 56$)	Case 4 ($N = 62$)
Time (29.5%)				
– Absolute ($N = 12$)	=	6	3	3
– Relative ($N = 44$)	10	3	15	16
Space (22.6%)				
– Absolute ($N = 29$)	7	5	3	14
– Relative ($N = 14$)	4	3	4	3
Time-Space (47.9%)				
– Absolute ($N = 31$)	9	=	10	12
– Relative ($N = 60$)	14	11	21	14

Findings show that the combination of space-time referentials (47.9%) by far
outweighs the single categories of temporal referentials (29.5%) and spatial ones
(22.6%), and that relative spatiotemporal referentials ($N = 60$; 65.9%) occur almost

---

6. Cf. Beatty (2008: 74).

twice as frequently as absolute spatiotemporal referentials ($N = 31$; 34.1%). This suggests that the orientation component of PENs is governed by the *inseparability* of space and time, and that it is coherently characterized by the prevalent occurrence of *relative* spatiotemporality of illness.

4.2   The evaluative functions of the PENs

Embedded evaluation, evaluative action and external evaluation are found to occur in all the sample PENs. Embedded evaluation is expressed through the Labovian syntactic categories of intensifiers, comparators and correlatives. Intensifiers are mainly represented through evaluative adverbs and adjectives:

(7) a.   I suffered <u>exceedingly</u> from the spasms; I have bine for this some time past <u>violantly</u> affected with a Nervous disorder; in this last simptom I was <u>gratly</u> afflicted with it [my breathing].   (Robert Ligertwood, 1777)

   b.   I fill the simptoms to be <u>shocking</u>; my cough was <u>very troblesome</u> I was sized in the Morning with such <u>dreadful</u> simptoms from a confusion in my head.   (Jane Webster, 1780)

In Example (7a.), evaluative adverbs (*exceedingly, violantly, gratly*) are used as amplifiers to enlarge the meaning of the verbs indicating the sufferer's feelings of pain (*suffered, affected, afflicted*); in Example (7b.), evaluative adjectives (*shocking, very troblesome, dreadful*) intensify the extent of the quality of the symptoms and cough. As these evaluative markers strengthen illness meanings, they contribute significantly to attracting the physician's attention to the centrality of these narrative units.

   On the other hand, comparators operate in a more elaborate manner through the combination of different evaluative markers, while correlatives connect the disorder with potential causal events:

(8) a.   I was advised to endeavour to reduce my corpulency, least my disorder <u>should</u> retturn again and be <u>more troblesome</u> to me if I continued <u>very</u> fat.   (Jane Webster, 1781)

   b.   but my stomach <u>not</u> being <u>so well</u>, owing probably to some <u>ill chosen</u> food.   (Anthony Chamier, 1780)

In Example (8a.), the narrator compares her current state of obesity to its consequences in the case of a disease relapse, expressing her evaluation based on reasonable medical advice. A complex evaluative comparison is construed through the combined use of the modal verb *should*, the graded quantifier *more*, the pejorative adjective *troblesome* and the intensifier *very* to mark the rather extreme likelihood that the narrator's health will worsen if she remains obese. In Example (8b.), instead, the state of the narrator's stomach is scaled down through the downtoner + adjective construct (*not so well*). Hence, this syntactic device first implicitly compares

the current state of disorder with a prior healthier condition and then correlates it with the possible causal event (*owing probably*) of an inappropriate food choice, marked by the negative evaluative qualifier (*ill chosen*).

Furthermore, markers of evaluative action particularly depict the dramatic sequence of illness events by highlighting sufferers' evaluative decisions about remedies:

(9)   I am <u>much</u> determined not to have recurrence to the absorbent, if it is <u>possible</u> to resist it but I must own I feel much despondence at present, as the <u>constant</u> pain interrupts my sleep and in <u>some</u> degree effects my spirits.

<div align="right">(Anthony Chamier, 1780)</div>

The painful event in Example (9) is marked by the negative evaluative adjective (*constant*) and by the moderate-degree quantifier (*some*) to indicate the narrator's stance towards his state of health. Stance is further shaped by the subjective decision to avoid the remedy of absorbing powder, marked by the intensifier *much* and in connection to the deontic possibility of resistance (*if it is possible*). Interestingly, the historical present tense (*am, is, feel, interrupts, effects*) seems to be used to heighten the dramatic force of the PEN, thereby creating narrative immediacy. In turn, this conveys a sense of urgency, which makes the pain experience more visible and persisting, besides helping elicit empathy from the physician.

Additionally, external evaluation appears to function as a premise to the main narrative since it is frequently placed in the initial narrative unit of the PENs:

(10)   a.   My Father, who was of your profession, was also <u>very</u> corpulent.

<div align="right">(Jane Webster, 1780)</div>

   b.   I shall always gratefully acknowledge my obligations to you for the <u>most judicious, candid, and accurate</u> observations which your Letter contains.

<div align="right">(Jane Webster, 1781)</div>

In Example (10a.), the narrator evaluates her father's body shape through the explicative intensifier (*very*) as a premise to her own health issue. Through this external evaluation, the reader is informed that the narrator's obesity appears to be a familial disease and that she has consulted Dr. Cullen despite having a doctor in her own family. In Example (10b.), instead, the narrator steps outside the story to directly give Cullen positive feedback about his counselling. This is accomplished through the use of a triple graded attributive (*most judicious, candid, and accurate*), which signals trust in the physician and positively predisposes the reader to the advice-seeking units of the narrative.

On the whole, 371 tokens of evaluative devices were recorded, suggesting that greater significance is attributed to evaluation compared to spatiotemporal referentials ($N = 190$) by all four narrators (see Table 2).

Table 2. Distribution of the Evaluative Functions of PENs per Evaluation Type

Types of evaluation ($N = 371$)	Case 1 ($N = 75$)	Case 2 ($N = 74$)	Case 3 ($N = 126$)	Case 4 ($N = 96$)
Embedded Evaluation ($N = 200$; 54%)	47 (62.6%)	37 (50%)	72 (57.2%)	44 (45.8%)
Evaluative Action ($N = 104$; 28%)	21 (28%)	17 (22.9%)	27 (21.4%)	39 (40.6%)
External Evaluation ($N = 67$; 18%)	7 (9.4%)	20 (27.1%)	27 (21.4%)	13 (13.6%)

The uneven distribution of the evaluative functions across all PENs shows that the occurrence of embedded evaluation (54%) is almost double that of evaluative action (28%) and significantly much higher than external evaluation (18%). Quantitative findings thus indicate that all of the patients mostly remain within the narratives to describe their feelings at the time of the illness events. However, they also mediate evaluative action regarding other possible remedies particularly in Case 4: e.g. "I get one of the usual size introduced as easily as ever […] a Bougie of the largest size (tho I do get it in) gives me rather more pain" (Henry Lochhead, 1777). In the Labovian sense, evaluative action is here constructed to heighten the narrative atmosphere. The first action is evaluated as a "typical" remedy through adjectival (usual) and adverbial (easily) markers; the variant action is meant to draw Cullen's attention to the increased degree (more) of pain experienced. It is, however, mitigated by the expression *tho I do get it in* and the pre-modifier *rather*, which pragmatically suggest a strategy of negotiation with Cullen.

Finally, the sparse use of external evaluation indicates that all of the narrators only occasionally interact with Dr. Cullen as their direct interlocutor, and stand outside the action mostly to provide information about their epistolary exchanges. One of them writes, for example, that "by the time I can receive your answer it will be the period to remove to Matlock" (Anthony Chamier 1780).

5.   Cullen's medical responses: Discourse strategies

A frequent use of *emotional talk* ($N = 60$) and *procedural talk* ($N = 53$) as *socio-emotional* discourse strategies is found to characterize Cullen's twelve epistolary responses. Emotional talk is mostly centred on offering *reassurance* to patients:

(11)   I am <u>certain</u> that the omission of these [absorbent powders] will allow the Rust of Steel to have more effect.   [Cullen to Chamier, 1780]

Cullen's response is shaped by "cognitive" rather than "affective" reassurance, whereby the epistemic stance marker (*certain*) indicates the physician's rather strong degree of certainty according to the basic medical knowledge available during the eighteenth century. In this regard, the reader may learn that metal-based compounds (*the Rust of Steel*) were widely used in the treatment of disease conditions even during the eighteenth century, although their therapeutic use was challenged by the toxic effects produced. Moreover, cognitive reassurance appears to have a systematic impact on changing patients' beliefs, increasing their understanding, besides yielding higher patient satisfaction (see Example 10b). Affective reassurance, instead, produces an immediate response in reducing concerns and therefore is not deemed to function owing to the epistolary nature of the medical interaction. Although "medical reassurance aims to reduce patients' illnesses", or "cognitions about disease", it does not, however, take "the psychological complexities" into account on its own (Coia and Morley 1998: 377). Cullen thus strategically handles the emotional components of illness by combining cognitive reassurance with "cognitive empathy", or "empathic accuracy", i.e., "more complete and accurate knowledge about the contents of another person's mind, including how that person feels" (Hodges and Myers 2007: 297):

(12)   It is possible that you may never be able to retain your urine so long as many others do or piss so slowly as you could do 20 years ago.

(Cullen to Lochhead, 1775)

In Example (12), *it is possible that* expresses knowledge-based possibilities concerning the patient's health problem, but also suggests Cullen's medical authority as the deontic source. The moderate degree of epistemic possibility allows Cullen to discursively construct cognitive empathy (*you may never be able; as you could do 20 years ago*), and thus indicates how he strategically interprets the patient's feelings about his urinary incontinence, and thereby legitimizes the narrative.

On the other hand, the scant use of affective empathy or "empathic concern" suggests that its functional purpose is mostly meant to respond to patients' narrative immediacy (see Example (9)) and, more in general, it seems to be an integral part of medical deontology:

(13)   I am heartily concerned to find that your nerves and stomach are not yet so firm.   (Cullen to Ligertwood, 1781)

The adverbial intensifier (*heartily*) in Example (13) marks the physician's "other-oriented value assessment […] that is congruent with the perceived welfare (well-being) of the other" (Batson 2011: 42) to express empathic concern. However, affective empathy also acquires the historical significance of securing public respect owing to the status held by physicians in the eighteenth century. As Wild (2006: 18–19)

clarifies, "[…] medicine as a profession had not achieved the prestige that it would in the nineteenth century. The eighteenth-century physician's practice and reputation depended entirely on the patronage and favour of his clients". Hence, Cullen is also found to express "a growth of professional feeling, which led to a struggle for improved status" (Hamilton 1951: 141) in spite of being a much sought-after consultant with an extensive epistolary practice (1755–1790), and a distinguished Faculty member of Edinburgh's medical school. In a letter addressed to the Lord Advocate Henry Dundas in 1786, Cullen admits "I trust I may be allowed to claim some merit with the town of Edinburgh and some right to their favour" (cited in Risse 1974: 338).

Regarding procedural talk as the other most prevalent strategy, responses show a constant use of concise medical instructions:

(14) a. Take half-an-ounce of powdered Peruvian bark[7] in a glass or marble Mortar; Nothing more proper than the cold bathing and the going on horseback as often as the weather allows.   (Cullen to Ligertwood, 1777)
      b. As slight ailments may go away with low living, watery drinks or by adding an Ounce of Gum Arabic[8] to the bottle of water.
(Cullen to Ligertwood, 1778)
      c. Moderation in diet, animal foods only once a day and then very moderately.
(Cullen to Webster, 1780)

On the whole, these instructions show the patient-centredness of medical advice as a result of Cullen's artful interpretation of PENs. These instances also shed light on the physician's practice of herbal medicine (Examples (14a., b.)), on the natural remedies he prescribed to recreate the Spirits according to humoral medicine (Example (14a.)), and his modern-sounding medical concerns about the issue of obesity (Example (14c.)). Yet, Cullen also appears to rely on the "symptom-based model of illness", whereby the symptom was the illness, thus ensuring that patient dominance was preserved throughout this period (Kaba and Sooriakumaran 2007):

(15) a. I believe I need not offer you any particular directions, your own experience will be a sufficient Guide while you have firmness and resolution to follow it; you must yourself judge; and hereafter take your measures accordingly.
(Cullen to Chamier, 1780)
      b. With regard to Diet I must think that your own experience has given you a great deal of instruction.   (Cullen to Ligertwood, 1781)

---

7. Bark of various species of the Cinchona tree, from which quinine was later isolated, and commonly used as a febrifuge (http://cullenproject.ac.uk/items/i1/), and particularly as the world's first anti-malaria drug.

8. Hardened sap of two species of the acacia tree (http://cullenproject.ac.uk/items/i75/).

Hence, the expressions *your own experience* and *you must yourself judge* explicitly confirm that "the sufferer habitually played an active and sometimes a decisive role in interpreting and managing his own state of health" (Porter 1985a: 138).

Overall, a total of 197 instances of Dr. Cullen's discourse strategies were recorded with a much higher frequency of socio-emotional strategies (81.73%) than task-focused ones (18.27%). Hence, Cullen's responses seem to denote his profound willingness to build doctor-patient relationships through his professional epistolary practice. This is mainly accomplished through emotional talk (30.45%), based on offering reassurance, and procedural talk (26.90%) providing medical instructions about treatment and remedies (see Table 3). This denotes a balance between Cullen's caring human interaction as a pillar of his service and the physician's socially-oriented mission of improving patients' physical health. Interestingly, this kind of interaction appears to be a forerunner of the modern medical construct of patient-centredness, which is seen as possibly leading to "improved healing relationships" (Greene et al. 2012: 50).

Table 3. Dr. Cullen's strategic responses (based on Roter and Larson 2002)

RIAS categories (N = 197)	Case 1 (N = 44)	Case 2 (N = 21)	Case 3 (N = 56)	Case 4 (N = 76)
Task-Focused Strategies (N = 36; 18.27%)				
Data-Gathering (N = 4; 2.03%)	–	–	4	–
Patient Education (medical info-giving) (N = 32; 16.24%)	8	6	7	11
Socio-emotional strategies (N = 161; 81.73%)				
Emotional Talk (N = 60; 30.45%)	– reassurance (7; 41.2%) – empathy (7; 41.2%) – concern (3; 17.6%)	– reassurance (1; 1.6%) – –	– reassurance (8; 61.5%) – empathy (4; 30.8%) – concern (1; 7.7%)	– reassurance (14; 48.3%) – empathy (10; 34.5%) – concern (5; 17.2%)
Negative Talk (disapproval/ criticism) (N = 16; 8.12%)	–	1	7	8
Patient Activation (N = 22; 11.17%)	6	4	5	7
Positive Talk (approval) (N = 4; 2.03%)	–	2	2	–
Procedural Talk (medical instructions) (N = 53; 26.90%)	12 (27.3%)	7 (33.3%)	17 (30.4%)	17 (22.4%)
Social Talk (nonmedical) (N = 6; 3.06%)	1	–	1	4

6. Concluding remarks

The epistolary cases examined in the present study have considered narrators with different social profiles, namely a political figure, an upper-class lady, a young intemperate man and a business traveller who were affected by a variety of health issues as reflected in their PENs.

On the whole, findings reveal that sufferers use spatiotemporal referentiality to mediate PENs as factual narratives, and engage in dynamic processes of frame switching by moving from absolute to relative referentials to foreground their illness accounts. However, the constant use of absolute referentials also shows how illness narrators bind themselves to their sociohistorical contexts, thus offering valuable insights into eighteenth-century medical practices. The narrators further exploit spatiotemporal referentiality for the functional purpose of monitoring their own illness trajectories, allowing the reader to gain an understanding of how the dynamic PEN structure of progression, stability and/or regression can be construed. Ultimately, these findings confirm that "the content of a text is not merely an enumeration of referents; an important part of the content is the relations that the text establishes between the referents" (Kallgren 1978: 150, cited in Brown and Yule 1983: 201).

Additionally, findings show how evaluation forms "part of the narrative which reveals the attitude of the narrator towards the narrative by emphasizing the relative importance of some narrative units" (Labov and Waletzky 1997: 32). Consistently, embedded evaluation is predominantly used to highlight the paramount importance of evaluating symptoms as illness principally through the Labovian syntactic categories of intensifiers and comparators and to a lesser extent through correlatives. Evaluative action is also employed to display the sequentiality of illness events and evaluate decisions about remedies. Although external evaluation is seldom used, its main functional purpose is to justify the actual epistolary narrative, besides allowing the narrator to step out of the account so as to directly acknowledge Cullen for his expertise.

As for the physician's responses, these show a significant use of emotional talk grounded in cognitive reassurance and cognitive empathy to ensure patient participation and interactional relations. The equally significant use of procedural talk featured by medical instructions emphasizes a more socially-oriented nature of medical advice. Consistently, the scant use of task-focused strategies in general, and of data gathering in particular, point out the importance ascribed to patient narratives under the influence of the Enlightenment symptom-based model which, by its nature, ensured active patient participation in medical interaction.

While the study has drawn on a small dataset, and a more extensive analysis is needed for further generalizations, it has, nonetheless, shown that the

eighteenth-century epistolary practice significantly represents "the voice of the life-world" (Mishler 1984), which has long been silenced by modern medicine. As "the sufferer's role in the history of healing […] has been routinely ignored by scholars" (Porter 1985b: 176), the discourse analysis has helped shed light on "the considerable agency of the sufferer" (Withey 2013: 124) in the epistolary representation of historical medical discourse.

In a diachronic perspective, the in-depth analysis of the linguistic resources deployed in PENs has ultimately underlined the importance of illness narratives. These are, in fact, currently regaining momentum "to provide to medicine what it lacks today", i.e., "the requirement for narrative knowledge and skills in the care of the sick" in order "to understand the plight of [patients] by participating in his or her story with complex skills of imagination, interpretation, and recognition" (Charon 2006: 8–9, 12). This renewed emphasis is witnessed by the recent development of the field of narrative-based medicine (NBM), which is "bringing the patient as a subject back into medicine" (Kalitzkus and Matthiessen 2009: 80–81). This also calls for further research on the historical evolution of PENs in the current e-Health era which fosters digital discourse practices of narrating illness in new interactive ways (Plastina 2015).

References

Batson, C. Daniel. 2011. *Altruism in Humans*. New York: Oxford University Press.

Beatty, Heather. 2008. "Quacks, Social Climbers, Social Critics, and Gentlemen Physicians: The Nerve Doctors of Late Eighteenth Century-Britain." In *The English Malady: Enabling and Disabling Fictions*, ed. by Glen Colburn, 67–94. Newcastle-upon-Tyne: Cambridge Scholars Publishing.

Brown, Gillian, and George Yule. 1983. *Discourse Analysis*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511805226

Charon, Rita. 2006. *Narrative Medicine: Honoring the Stories of Illness*. Oxford: Oxford University Press.

Coia, Pete, and Stephen Morley. 1998. "Medical Reassurance and Patients' Responses." *Journal of Psychosomatic Research* 45 (5): 377–386.

De Fina, Anna, and Barbara Johnstone. 2015. "Discourse Analysis and Narrative." In *The Handbook of Discourse Analysis* (2nd edition), ed. by Deborah Tannen, Heidi E. Hamilton, and Deborah Schiffrin, 152–167. Malden, MA: Wiley Blackwell.

Emmott, Catherine. 1997. *Narrative Comprehension: A Discourse Perspective*. Oxford: Oxford University Press.

Fairclough, Norman. 1992. *Discourse and Social Change*. Cambridge: Polity Press.

Gergen, Kenneth J., and Mary M. Gergen. 1986. "Narrative Form and the Construction of Psychological Science." In *Narrative Psychology: The Storied Nature of Human Conduct*, ed. by Theodore R. Sarbin, 22–44. New York: Praeger.

Greene, Sarah M., Leah Tuzzio, and Dan Cherkin. 2012. "A Framework for Making Patient-Centered Care Front and Center." *The Permanente Journal* 16 (3): 49–53. https://doi.org/10.7812/TPP/12-025

Hamilton, Bernice. 1951. "The Medical Profession in the Eighteenth Century." *The Economic History Review* 4 (2): 141–169. https://doi.org/10.2307/2599120

Hawkins Hunsaker, Anne. 1999. *Reconstructing Illness: Studies in Pathography*. West Lafayette, IN: Purdue University Press.

Hodges, Sarah D., and Michael W. Myers. 2007. "Empathy." In *Encyclopedia of Social Psychology*, ed. by Roy F. Baumeister, and Kathleen D. Vohs, 296–298. Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781412956253.n179

Hydén, Lars-Christer. 1997. "Illness and Narrative." *Sociology of Health & Illness* 19 (1): 48–69. https://doi.org/10.1111/j.1467-9566.1997.tb00015.x

Jurecic, Ann. 2012. *Illness as Narrative*. Pittsburgh: University of Pittsburgh Press. https://doi.org/10.2307/j.ctt5hjr8p

Kaba, Rie, and Prasanna Sooriakumaran. 2007. "The Evolution of the Doctor-Patient Relationship." *International Journal of Surgery* 5 (1): 57–65. https://doi.org/10.1016/j.ijsu.2006.01.005

Kalitzkus, Vera, and Peter Matthiessen. 2009. "Narrative-Based Medicine: Potential, Pitfalls, and Practice." *The Permanente Journal* 13 (1): 80–86. https://doi.org/10.7812/TPP/08-043

Labov, William. 1972. "The Transformation of Experience in Narrative Syntax." In *Language in the Inner City*, ed. by William Labov, 354–396. Philadelphia, PA: University of Pennsylvania Press.

Labov, William, and Joshua Waletzky. 1997. "Narrative Analysis: Oral Versions of Personal Experience." *Journal of Narrative and Life History* 7 (1–4): 3–38. https://doi.org/10.1075/jnlh.7.02nar

Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511813313

Louis-Courvoisier, Micheline, and Alexandre Mauron. 2002. "'He found me very well; for me, I was still feeling sick': The Strange Worlds of Physicians and Patients in the 18th and 21st centuries." *Journal of Medical Ethics: Medical Humanities* 28 (1): 9–13.

Louis-Courvoisier, Micheline, and Séverine Pilloud. 2004. "Consulting by Letter in the Eighteenth Century: Mediating the Patient's View." In *Cultural Approaches to the History of Medicine: Mediating Medicine in Early Modern and Modern Europe*, ed. by Willem de Blécourt, and Cornelie Usborne, 71–88. Basingstoke and New York: Palgrave Macmillan. https://doi.org/10.1057/9780230287594_5

McEnery, Tony, Richard Xiao, and Yukio Tono. 2006. *Corpus-Based Language Studies: An Advanced Resource Book*. London and New York: Routledge.

Mishler, Elliot G. 1984. *The Discourse of Medicine: Dialectics of Medical Interviews*. Norwood, NJ: Ablex.

Overton, Bill. 2007. *The Eighteenth-Century British Verse Epistle*. Basingstoke: Palgrave Macmillan. https://doi.org/10.1057/9780230593466

Plastina, Anna Franca. 2015. "The Discourse of Comforting: The Case of Online Health Support Groups." *Token: A Journal of English Linguistics* 4: 37–56.

Porter, Dorothy, and Roy Porter. 1989. *Patient's Progress: Doctors and Doctoring in Eighteenth-Century England*. Stanford, CA: Stanford University Press.

Porter, Roy. 1985a. "Lay Medical Knowledge in the Eighteenth Century: The Evidence of *The Gentleman's Magazine*." *Medical History* 29 (2): 138–168. https://doi.org/10.1017/S0025727300043970

Porter, Roy. 1985b. "The Patient's View: Doing Medical History from Below." *Theory and Society* 14 (2): 175–198. https://doi.org/10.1007/BF00157532

Risse, Guenter B. 1974. "'Doctor William Cullen, Physician, Edinburgh': A Consultation Practice in the Eighteenth Century." *Bulletin of the History of Medicine* 48 (3): 338–351.

Roter, Debra, and Susan Larson. 2002. "The Roter Interaction Analysis System (RIAS): Utility and Flexibility for Analysis of Medical Interactions." *Patient Education and Counseling* 46 (4): 243–251. https://doi.org/10.1016/S0738-3991(02)00012-5

Schaeffer, Jean-Marie. 2014. "Fictional vs. Factual Narration." In *Handbook of Narratology*, Vol. 1, ed. by Peter Hühn, Jan C. Meister, John Pier and Wolf Schmid, 179–196. Berlin and Boston: De Gruyter. https://doi.org/10.1515/9783110316469.179

Steinke, Hubert, and Martin Stuber. 2004. "Medical Correspondence in Early Modern Europe: An Introduction." *Gesnerus* 61: 139–160. https://doi.org/10.24894/Gesn-en.2004.61009

Stolberg, Michael. 2013. "Examining the Body, c. 1500–1750." In *The Routledge History of Sex and the Body: 1500 to the Present*, ed. by Sarah Toulalan, and Kate Fisher, 91–105. London and New York: Routledge. https://doi.org/10.4324/9780203436868-ch-5

Wild, Wayne. 2000. "Doctor-Patient Correspondence in Eighteenth-Century Britain: A Change in Rhetoric and Relationship." *Studies in Eighteenth-Century Culture* 29: 47–64. https://doi.org/10.1353/sec.2010.0288

Wild, Wayne. 2006. *Medicine-by-Post: The Changing Voice of Illness in Eighteenth-Century British Consultation Letters and Literature*. Amsterdam and New York: Rodopi. https://doi.org/10.1163/9789401202350

Withey, Alun. 2013. *Physick and the Family: Health, Medicine and Care in Wales, 1600–1750*. Manchester: Manchester University Press.

Communicating authority

Self-mentions in Early Modern English medical narratives (1500–1700)

Karoliina Ollikainen

University of Glasgow

This sociopragmatic case study examines 31 narrative passages across 23 different early modern medical texts in order to determine how medical writers use the narrative discourse mode to construct their own authority as medical professionals. Based on how the authors present themselves through self-mentions, the narratives have been divided into two different types: treatment narratives and narratives of discovery. The findings of the study suggest that early modern medical writers take on different authoritative roles in these narratives based on the aims of their text. Overall, however, all of the narratives are used to highlight the medical practitioners' professional skill and experience, which adds to the authoritativeness of their claims.

Keywords: Early Modern English medical texts, empiricism, narrative discourse, historical pragmatics

1. Introduction

The early modern period was a time of scientific change and uncertainty, which also affected what was considered to be safe, authoritative medical knowledge. Prior to the latter half of the sixteenth century, by contrast, learned medicine had been relatively stable and clearly defined. It was based on the teachings of trusted, irrefutable authorities, namely the Ancient Greeks, such as Hippocrates (c. 370–460 BC) and Galen (130–210 AD), as well as influential medieval Arabic writers, such as Muhammad Ibn Zakariya al-Razi (854–925/932), Ibn Sīnā (980–1037), and Ibn Zuhr (1094–1162), known in the West by their Latinised names Rhazes, Avicenna, and Avenzoar. As French (2003: 81–82) notes, the authority of the university-educated doctor, therefore, was based on their familiarity with these authoritative sources and their ability to interpret and apply their teachings. This way of learning and

generating authoritative knowledge can be referred to as the scholastic method of enquiry, which focuses on comparing and reconciling the views of different important authorities (Siraisi 1990: 73, 76). When it came to medical texts, the writers could prove that they were knowledgeable and reliable by referencing the right external sources and demonstrating that their views aligned with the writings of the key authorities.

Over the course of the early modern period, however, European approaches to science gradually shifted away from the scholastic tradition and moved towards empiricism. The decline of the scholastic worldview also undermined the authority of traditional scholastic sources of information, and by the end of the seventeenth century, the top-levels of science had largely moved away from the teachings of the Ancient Greeks.[1] After the long-lasting tradition of scholastic medicine, this was a drastic change. Furthermore, no new theory unified the field of medicine like Hippocratic medicine had done in the previous centuries, but rather, early modern medicine was characterised by a number of different branches competing with each other (Wear 2000: 360). The writers of medical texts, therefore, needed to find different ways of justifying their arguments, and practitioners had to construct their authority as medical professionals in new ways.

This sociopragmatic case study will examine how medical writers navigated these changes in medical discourse by analysing self-mentions in narrative passages. Based on how these self-mentions are used within their textual context, I argue that the pragmatic function of narratives told in 1st person singular is twofold. First, they play an integral role in effectively communicating medical knowledge to the reader. The narrative passages allow the medical writers to discuss the progression of illnesses and the course of treatment in a great amount of detail, and the concrete examples they provide make it easier for the reader to potentially replicate their methods. Second, 1st person narratives are also inherently autobiographical. As has been argued by Hyland and Tse (2004: 169), self-mentions are an interactional metadiscursive feature that makes the writer's presence in the text explicit. Therefore, they give us information about how the authors want to present themselves to their readers, which is also the focus of the present study. More specifically, the research question examined here is how medical writers use the narrative discourse mode as a vehicle for talking about their medical knowledge and professional experience. Through examining self-mentions in narratives, my aim is to establish how early modern medical writers construct their professional identity and how they communicate that authoritative position to their readers.

---

1.   It should be noted, however, that the everyday practice of medicine changed relatively little during this time, and Hippocratic methods such as bloodletting and purging continued to be commonly used (Wear 2000: 434).

2.   Material and methods

This case study examines 31 narrative passages across 23 different early modern medical texts. These medical texts come from the *Early Modern English Medical Texts* corpus (EMEMT, Taavitsainen et al. 2010), which covers the time period between 1500 and 1700. This corpus consists of about two million words, and its material is divided into six subcategories based on the topics the texts cover as well as the genres of medical writing they fit into.

The narratives included in this study have been collected as part of a larger corpus linguistic project on authority in early English medical writing (cf. Ollikainen 2019). The passages relevant to the project were located using the mental verbs BELIEVE, THINK, KNOW, OBSERVE, and PERCEIVE, as well as the communication verb SAY. These verbs were chosen based on the findings of previous studies on early English medical writing. The five mental verbs tend to be connected to discussing subjective knowledge and empirical observations (Taavitsainen 2002: 207), which makes them particularly useful for locating passages related to empirical argumentation. The communication verb SAY, on the other hand, not only serves important narrative and metatextual functions (Jajdelska 2018) but is also associated with the authoritative scholastic style (Pahta and Taavitsainen 1998: 174–175), which makes SAY useful for identifying passages where authors position themselves as the authoritative source of information. The self-mentions captured by this approach were divided into two main categories based on how the authors justify their knowledge: *experience* and *book learning* (Ollikainen 2019: 211). When writers justify their authority with *book learning*, they emphasise how well-read and learned they are by referring their readers to authorities that agree with them. *Experience*, by contrast, was used to refer to instances where authors base their claims on observations they have made or talk about their practical knowledge, for instance by mentioning patients they have successfully treated before. The present case study focuses on the latter and examines the different ways in which the medical writers discuss their past experiences in narrative passages.

Overall, the corpus study captured 1,426 self-mentions, whereas 1st person singular pronouns (*I, me, my, mine*) have 4,789 concordance hits in EMEMT. This means that the corpus search captured approximately 30% of all self-mentions in EMEMT. Out of these self-mentions, 56 occur in narrative passages, which all in all amount to 31 unique narratives. Without doubt, there are more 1st person singular narratives in EMEMT than the 31 examined here, and determining their proportion in the whole corpus would require a different sampling strategy. However, what kind of information the authors prioritise and the details they choose to include in these narratives can demonstrate how they want to present themselves to the reader. Furthermore, the commonalities between different texts by different authors

provide a basis for identifying the functions served by these narratives, as their shared features will highlight the ways in which early modern medical authors tended to construct their identity as medical professionals.

What counts as a narrative passage is defined here in terms of the narrative discourse mode, i.e. they are passages that describe how events progressed as the narrative time advances (Smith 2003: 14–16). Discourse modes are a valuable tool for analysing the local structure of texts as they make it possible to, for example, identify short narrative passages embedded within texts that are not otherwise told in the narrative format. While various suggested models for discourse modes have fallen in and out of use since their conception in the nineteenth century (Connors 1981: 444–446), the concept has been especially clarified by Smith's (2003) classification which in addition to narratives recognises four other discourse modes: *information, description, report*, and *argument*. While this classification partially overlaps with the widely recognised five text types (c.f. Werlich 1976), Smith's approach makes a useful distinction between different types of temporal and atemporal modes of discourse. According to Smith's definition, *information* and *description* are both discourse modes that aim to communicate factual information to the reader. Whereas informative passages are atemporal and focus on generalised or universal facts (Smith 2003: 17), descriptions give information about a specific event or a person (Smith 2003: 28–29). However, even though descriptions can give information about specific events that took place at a specific point in time, they are static and do not discuss in what order events occur and how they progress. The passages that have dynamic temporality and describe the order in which events progress through time Smith categorises either as a narrative or a report (2003: 14, 16). Although reports and narratives both detail how particular events transpired, Smith's distinction is that reports, such as live news stories, are written down in real time (2003: 16). Due to the nature of my material, all of these passages have been categorised as narratives, as all of the texts have been written down after the fact.

As mentioned above, the present study includes 31 narrative passages that come from 23 different early modern medical texts. The corpus texts included in the study are listed in Table 1. Eleven of the narratives examined in this case study are from the sixteenth century, and the remaining 20 come from seventeenth century texts. When it comes to the categories in the corpus, surgical texts are the most represented in the sample with 13 narratives. The second most common source of 1st person singular narratives is the Royal Society's journal *Philosophical Transactions* with nine narratives. The most well-represented author is William Clowes (1543–1604), whose *Booke of Obseruations* (1596), *Lves Venerea* (1596), and *Cure of Struma* (1602) collectively account for six narratives. The fact that Clowes's narratives make up roughly 19% of the material is potentially problematic, as some features identified by the analysis could be representative of Clowes's individual

Table 1. Texts included in the study

Corpus category in EMEMT	Author	Text
GENERAL TREATISES AND TEXTBOOKS	Robert Couch	*Praxis Catholica* (1680)
THE *PHILOSOPHICAL TRANSACTIONS*	Jean-Baptiste Denys	Vol. 2 (32):617–23 (1668)
		Vol. 3 (36):710–5 (1668)
	Henry Stubbes	Vol. 3 (36):699–709 (1668)
	Thomas Platt	Vol. 7 (87):5060–7 (1672)
	Robert Boyle	Vol. 7 (89):5108–16 (1672)
	Edmond King	Vol. 16 (185):228–231 (1686)
	Allan Mullen	Vol. 16 (192):486–8 (1688)
	Robert Boyle	Vol. 17 (197):627–41 (1693)
	Daniel Turner	Vol. 18 (207):24–25 (1694)
RECIPE COLLECTIONS AND *MATERIA MEDICA*	Nicolás Monardes	*Ioyfvll Newes* (1580)
	Giambattista della Porta	*Natural Magick* (1658)
TEXTS ON SPECIFIC DISEASES	John Caius	*Against Sweatyng Sicknesse* (1552)
	William Clowes	*Lves Venerea* (1596)
		Cure of Struma (1602)
	William Cockburn	*Continuation of the Account of Distempers* (1697)
TEXTS ON SPECIFIC METHODS	John Bartlet	*Warming Stone* (1640)
SURGICAL AND ANATOMICAL TREATISES	William Bullein	*Bulleins Bulwarke* (1562)
	Thomas Gale	*Institution of a Chirurgian* (1563)
	William Clowes	*Booke of Obseruations* (1596)
	Ambroise Paré	*Workes of Ambrose Parey* (1634)
	William Harvey	*Anatomical Exercises* (1653)
	John Colbatch	*Novum Lumen Chirurgicum* (1698)

writing style rather than medical writing in general. In addition to Clowes, the corpus results also include multiple narratives from Robert Boyle (2 passages), John Caius (2), John Colbatch (2), Thomas Gale (2), and Ambroise Paré (3). Therefore, in analysing these narratives, I have only focused on features that are shared by multiple narratives from different authors. Any features that occur in only a single narrative or only in narratives by one author have been disregarded.

My analysis of these narratives is based on the theoretical framework of historical corpus pragmatics and sociopragmatics. Historical pragmatics sees written texts as communicative acts that are shaped by the participants and the context within which they were produced (Jucker 1995: 9–10; Taavitsainen and Fitzmaurice 2007: 14). The focus of historical sociopragmatics, therefore, is to identify how the communicative choices made by the writer influence the meaning of the text and how the readers of that text deconstruct the writer's message (Taavitsainen and

Fitzmaurice 2007: 25). As the research questions explored by historical pragmatics depend on empirical data of real language use, corpus linguistic methods are used widely to systematically gather information about past communicative practices (Taavitsainen 2018: 527). The quantitative methods of corpus linguistics can help establish what kind of practices are common in a certain historical genre or during a given time period, which in turn can also indicate what kind of language use stands out as unusual (Taavitsainen and Fitzmaurice 2007: 29). However, when using corpus linguistic methods in pragmatics, decontextualisation can present a problem. As pragmatics is the study of language use in context, isolating individual instances of linguistic features from the rest of the text can lead to misrepresentations of the actual pragmatic functions served by those features (Taavitsainen 2018: 542). Therefore, corpus linguistic studies with a sociopragmatic focus need to be complimented by qualitative analysis of the source material in order to accurately interpret how the corpus results are used in context (see also Hiltunen, this volume).

Consequently, this case study takes the data collected for the larger, quantitative corpus project (Ollikainen 2019) and focuses on the qualitative analysis of the 56 narrative self-mentions captured by the corpus study. By using the methods of historical sociopragmatics and studying these self-mentions within their textual and historical context, the aim of the analysis is to determine how the medical writers choose to present themselves to their readers and what kind of professional identities they assume based on the kind of information the writers prioritise in their description of the events as well as what kind of details their narratives emphasise. These details include the role of the author and other participants in the narrative, the specifics of the patient's condition or other object of study, and the outcome of any treatments or experiments. Although it is beyond the scope of this study to determine to what extent the details of each narrative discussed here are true, and corroborating or disproving them is not likely to be possible, they demonstrate how medical writers choose to present themselves through 1st person narratives. In my analysis of these narrative passages, I have identified two distinct types of narratives that reflect different kind of authoritative roles: treatment narratives and narratives of discovery.

Generally speaking, these two types of narratives can fit under two broad genres of medical and scientific writing: case histories and experimental accounts. Although there are a few examples of illness narratives in late medieval surgical texts (Turner 2016: 62), case histories are mainly a feature of early modern medical writing. As has been discussed by Pomata (2010: 196–7), the emergence of medical case histories over the course of the sixteenth century highlights the increasing interest in medical practice over medical theory, as well as the focus on experience and observation over the writings of previous authorities. Case histories, therefore,

came to be used to legitimise medical practices, as they demonstrated the success of the methods used by the doctor (Pomata 2010: 207). Experimental accounts, on the other hand, became an established genre of scientific writing during the latter half of the seventeenth century, and the conventions of the genre were largely shaped by the Royal Society's journal *Philosophical Transactions*, first published in 1665 (Gotti 2011: 204, 209; see also Menzel, this volume). Experimental accounts are used to share the writer's observations with the reader, and the focus tends to be on describing how the writer has examined their object of study and what they have been able to perceive through their senses (Gotti 2011: 212–213).

However, even though what this study calls treatment narratives can roughly correspond to the genre of case histories, and experimental accounts are likely to include narratives of discovery, this is not necessarily always the case. What has been classified here as a treatment narrative could be a passage in an experimental account, and the same goes for narratives of discovery and case histories. Therefore, rather than focusing on the genre conventions the text fits into overall, the individual narrative passages studied here have been divided into two categories based on how the authors' use self-mentions to construct their professional identity.

3.   Treatment narratives

As the name suggests, treatment narratives are passages that prioritise information relating to treating a patient. These narratives focus on individual cases the doctor has treated and describe how a certain illness affected an individual patient. They also tend to give the reader information about established medical practices, such as how to diagnose and treat certain diseases, and detail how the illness progressed. Furthermore, they often include an element of self-promotion, as they showcase the author's professional skill by emphasising how effectively they are able to treat difficult and complicated cases. Overall, these treatment narratives tend to have a number of features in common: they highlight the seriousness of the patient's condition; include details that add to the credibility and verifiability of the narrative; and focus on either the effectiveness of the treatment or the accuracy of the practitioner's prognosis.

First of all, when it comes to emphasising the seriousness of the patient's condition, the narratives frequently focus on how long the patient had suffered from their ailment before finally being treated by the author. For instance, in *Praxis Catholica* (1680), Robert Couch tells that a maid he had in New England had previously struggled to find employment because "she was violently troubled with the Falling Sickness". Couch reports that he was immediately able to help her overcome her epilepsy

by administering a powder of his own design, but he also makes it clear that prior to this treatment the epilepsy had seriously impacted the patient's quality of life:

> (1)   […] I gave only one dose of this Powder, which freed her from her Fits, though **she had them before almost continually, day and night**, that whosoever had her, kept one to watch her, for **she would be ready to destroy her self**; I kept her about a quarter of a year, and she never had a Fit that I know of, I perswaded her to take another Dose two or three days after, but she would not; she told me she needed not, for she was confident she was perfectly cured. Since I hear she is married some where in the Country, and whether she had them since, I know not.   (Couch, *Praxis Catholica*, 1680)

Following this endorsement, Couch gives the recipe for his powder, making it possible for the reader to replicate the remedy. Based on the title page, this seems to be the explicit aim of Couch's text, as it promises that "[w]hereby any one of an ordinary Capacity may apprehend the true Cause of his Distempters, wherein his Cure consists, and the Means to effect it" (1680). *Praxis Catholica* includes a number of remedies for different ailments, and several of them are accompanied by similar accounts of how Couch has used them to successfully cure someone. These narratives, therefore, are provided as proof of how effective Couch's methods are. Consequently, they also suggest that Couch has a wealth of experience as a medical practitioner. This adds to the authoritativeness of Couch's text, because he not only teaches the reader how to treat certain ailments but also demonstrates through these narratives that he has an extensive track record when it comes to curing difficult medical conditions.

In a similar vein, when highlighting the seriousness of the patient's condition, a number of narratives also mention that other medical practitioners have previously tried and failed to help the patient. For instance, in William Clowes's *Booke of Obseruations* (1596), Clowes explains that he was specifically called to help a soldier whose bullet wound no one else had been able heal because the bullet remained unusually deep within the wound:

> (2)   In the yeere of our Lord 1573. I was sent for, vnto my singular good friend Master Richard Yoong, one of hir Maiesties Iustices of peace for Middlesex, who did earnestly request me, that I would for his sake, cure and heale, if it were possible, the aforesaid soldier, called M. Giles, for that he was knowen to be a very valiant man, which cure to performe, seemed to me very hard and difficult, for that he had been for the space of three yeeres, with **diuers very good Surgeons, both beyond the seas as also in England**, and yet his griefe did still reuerse, and breake out againe: the reason I perceiued was, for that the place, where the bullet lay, could neither by probation, nor coniecture be certainly knowen, and that was the chiefest cause I suppose, **why they failed in this cure**[.]   (Clowes, *Booke of Obseruations*, 1596)

Explicitly stating that other "very good surgeons" have not been able to help this soldier makes Clowes's eventual success even more impressive. According to the narrative, after encountering difficulties himself, Clowes manages to locate the bullet by administering various remedies to the wound that over the course of several days remove the scarring that was obscuring the bullet's location. Like Couch in *Praxis Catholica* (1680), Clowes also includes the recipes for each of these remedies and instructions for how to administer them. By including these details, Clowes presents himself as a skilful surgeon capable of handling a complicated case and, furthermore, that he is more knowledgeable than most other surgeons or, at the very least, knows more than the surgeons who had tried to help the patient before.

When Clowes finally successfully removes the bullet, he notes that "I called others in presence when I made [the] incision" (1596). This highlights another common feature in these narratives, which is mentioning that the events described by the author were witnessed by others. The fact that the accuracy of the narrative could potentially be verified by the other people present adds to the credibility of the author's claims. As has also been noted by Hiltunen (2010) in reference to articles published in the *Philosophical Transactions*, writers tend to emphasise the presence of eyewitnesses especially when they are of high status. John Colbatch's *Novum Lumen Chirurgicum* (1698), for example, explains how Colbatch effectively treated a wounded soldier, but the majority of his narrative is actually dedicated to demonstrating that a number of high-ranking people were involved in the process:

> (3) ONE Paterson a Scots-man, in **the Regiment of the Right Honorable the Lord George Hamilton (now Earl of Orkney)** was about 8 a clock one Evening run into the Belly about two Inches above the Navel on the right Side; the Sword[.] […] **By his Lordship's Command** I was sent for to take care of him[.] […] it was but two days before that **his Majesty** was graciously pleased by public Orders throout the Camp, that notice should be given to my most worthy Friend **Sir Henry Bellasis**, of the first Man wounded, that he might see whether I were capable of performing what I pretended to. […] On the fourth day he was able to go with me to Court, where **the General and other Officers** saw him very well. Besides the foresaid Wound, he was also wounded in two other places, and much bruised, of all which he perfectly recover'd.
>
> (Colbatch, *Novum Lumen Chirurgicum*, 1698)

Novum Lumen Chrirurgicum was first published in 1695, and the text describes the various injuries he treated during his time with the army in 1694 (Cook 2008). According to Cook (2008), at the time Colbatch was a controversial figure within the medical community in London. Colbatch was famous for his two remedies, "Vulnerary Powder" and "Tincture of the Sulphur of Venus", which he said were empirically proven to heal wounds faster than methods commonly used by surgeons (Cook 2008). These claims were received with scepticism, especially after

physicians failed to replicate his results when testing the powders on patients at the St. Bartholomew Hospital (*ibid.*). Colbatch's response to the failed trials was that his remedies had not been applied properly and with the help of a patron secured an opportunity to prove his skill by joining the military campaign in Flanders (*ibid.*). Therefore, *Novum Lumen Chrirurgicum* consists of numerous accounts of successfully treating severely wounded soldiers, and Colbatch states that the aim of his treatise is to prove that the wound's natural healing process is hindered "by the common Methods and Medicins of Chirurgeons", while his methods encourage healing (1698). Since Colbatch's time with the army and the writing of *Novum Lumen Chrirurgicum* were both motivated by trying to demonstrate the superiority of his methods, it is likely that Colbatch is mentioning the names of these prominent people in his narrative in order to prove that his success story can be verified by several reliable witnesses.

In addition to naming the people present during the events, the writers also scaffold their narratives with other details that make them potentially more verifiable. These details mainly boil down to when and where the narrative took place. The majority of the narratives mention at least one, if not both (Examples 4–6):

(4)   […] and now here I giue you to vnderstand that in **the yeere 1591. the xii. of Ianuary**, there was sent vnto me from Captaine Fleming, a certaine soldier which had a fistulas vlcer in his thigh […]
        (Clowes, *Booke of Obseruations*, 1596)

(5)   It fortuned that **in London .1562.** that there was a fraye made, and the one was thurste throughe the breast vnder the pappe, and out vnder the lefte shoulder blade so that of force the longes were persed through.
        (Gale, *Institution of a Chirurgian*, 1563)

(6)   Moreover, I will heere shew to the readers the places where I have had meanes to learne the Art of Chirurgery, for the better instructing of the young Chirurgion: and first in **the yeere 1536**, the great King Francis sent a great Army to **Thurin**, to recover the Cittyes and Castles, which the Marquesse of Guash, Lievtenant generall of the Emperor had taken[.]   (Paré, *Workes of Ambrose Parey*, 1634)

By providing these specifics, the writers are, consciously or not, providing potential evidence to the reader that the events they describe actually happened. Specifying the time and location can root the narrative in reality and make it easier for the reader to believe that the story is true, which adds to the authoritativeness of the writer's claims.

While the majority of the treatment narratives are success stories, the medical writers have also chosen, interestingly enough, to discuss cases in which the treatment fails. Within the sample of 31 narratives studied here, five include failed

treatments, and all but one end with the patient dying. Again, like the success stories discussed above, these narratives tend to emphasise the seriousness of the patient's condition. For instance, in Ambroise Paré's autobiographical work *Les oeuvres de M. Ambroise Paré* (1575), translated into English in 1634 by George Baker, Paré tells of a soldier who is wounded with a halberd for cheating in a game of dice, and Paré notes how remarkably deep the wound is and remarks that he could tell that the soldier would not survive:

(7) I saw one thing of great remark, which is this: that a souldier in my presence gave to one of his fellowes a stroake with an Halbard upon the head, **penetrating even to the left ventricle of the braine**, without falling to the ground. Hee that strooke him said, he had heard that he had cheated at Dice, and that he had drawne a great summe of money, and that it was his custome to cheat; I was called to dresse him, which I did as it were for the last, **knowing well that he would quickly die**: having drest him he returned all alone to his lodging, which was at least two hundred paces distant: I bid one of his companions send for a Priest to dispose of the affaires of his soule[.]

(Paré, *Workes of Ambrose Parey*, 1634)

As Paré predicted, the soldier dies after three days, and Paré explains that he decided to share this story because it is "a monstrous thing, that the Souldier fell not to the ground when he had received this great stroake, and was in good senses even till death" (1634). The explicitly stated aim of the narrative, therefore, is to give the reader information about serious head injuries. However, as Paré describes the last three days of the soldier's life, he reiterates multiple times that he knew that his administrations were ultimately futile. Consequently, the narrative also communicates to the reader that the soldier did not die due to Paré's lack of skill. Rather, being able to discern early on that the soldier's wounds were fatal is presented as testament to how knowledgeable and experienced Paré is as a surgeon.

Other narratives that describe failed treatments follow a similar pattern and focus on the importance of accurate prognosis. This is most evident in William Bullein's surgical text *Bulleins Bulwarke* (1562) where he first discusses what Hippocrates says about the danger of muscle cramps in wounded areas and how the possibility of cramps can complicate prognosis. To give an example of such a situation, Bullein follows this discussion with a narrative about a patient he helped treat when he was still an apprentice to another surgeon. Bullein mistakenly promises the patient that he will survive, and the surgeon he is working under takes him aside to teach him how to better estimate the likelihood of fatal convulsions:

(8)   […] **whiche thing trulie had, befallen to my self, when I promised a man his life, had not I been admonished of by maister Rasis**, who secretly and wittely, vnknowen to all the family, pluckte me by the elbowe, and warned me, with these, or like wordes in effecte. Consider with thy self, saieth he, as thou doest lightlie alwaie, the signes, which induce conuulsion in woundes. Where as there appereth no swelling in the sore, in good faithe, **I conceiue an euill opinion, in the pacientes life, whiche I haue gotten me, with long obseruacion**. For the matter, whiche should come forthe, to the pained place, is supped vp of the sinewes. Wherefore, I see a conuulsion euen at hande […].

(Bullein, *Bulleins Bulwarke*, 1562)

Master Rasis's prognosis turns out to be accurate, and the patient dies within five days. In this narrative, Bullein not only lists the signs a surgeon should look out for in this situation but also impresses upon the reader how important the ability to accurately predict the outcome is. Bullein finishes this cautionary tale by directly addressing the reader:

(9)   […] I appeale here to your secrete hartes, and bid you be ware, that **ye disgrace not your self thorowe rashenes**, but be ripe in prognosticacions, and circu~specte in obseruing of the times, of your workinges bothe for one helpe & others.   (Bullein, *Bulleins Bulwarke*, 1562)

As is evident from Bullein's warning, accurate prognosis is treated as a matter of protecting the practitioner's reputation, which is a line of thinking that can be traced back to Hippocrates' aphorisms (French 2003: 11–12). Therefore, narratives like these are not actually presented as failures even if the treatment is not successful. The medical practitioners involved are still trustworthy authorities in these stories, but instead of being able to save the patient, their skill, knowledge, and expertise go towards successfully predicting the outcome.

However, careful prognosis is not the only way the medical writers protect their professional reputation. In two of the narratives investigated in this case study, the writers turn their attention to the patient's behaviour instead and emphasise that the patient interfered with the course of the treatment. In *Cure of a Struma* (1602), William Clowes details the actions of a woman whose leg ulcer he treated together with a physician. Clowes prepares an ointment for the ulcer, and the physician prescribes a purgative to counteract the humors he believes to be the root cause. These remedies end up making the patient extremely sick, but Clowes explains that this is because she accidentally spread the internal medicine onto her leg and swallowed the external medicine:

(10) But note what ill hath followed by **the Maydes carelessness, and too much negligence**: In the morning after she came home to Yalding by 7. Of the clocke, and tooke one of the 3. Purgations which the Doctor gaue vnto her to take, shee set it vpon a stoole by the fire, where shee meant first to dresse her legge: in conclusion (**by great ouersight**) she laide the Purgation to her legge, and did eate vp the whole boxe of Aegiptiacum, which was nere 2. Oun. And (as she said afterwards) it was very vntoothsome and loathsome in tasting.

(Clowes, *Cure of Struma*, 1602)

As can be seen in this passage, the patient is characterised as someone who is too inattentive and negligent to follow the careful instructions Clowes tells he and the physician gave her, and therefore Clowes and his colleague are not responsible for the unfortunate end result of their treatments. Clowes goes on to report that it took the patient weeks to recover from the after effects of her mistake, and although the mix-up was eventually cleared up, the patient's father and another physician the family had consulted were originally under the impression that she had been given dangerous remedies by unskilled practitioners. Therefore, while Clowes's narrative provides detailed information on how the side effects were treated and how they helped her recover, the narrative also appears to have an element of damage control and seems to have been included in the text to restore Clowes's and the physician's reputation.

The second example of criticising the patient's actions comes from Jean-Baptiste Denys's letter published in the *Philosophical Transactions*. The letter concerns Anthony Mauroy, a French man who was the first human to receive a blood transfusion (Myhre 1990: 359). Denys and his associates gave the man calf's blood in order to help him with mental instability, and in his initial report to the Royal Society, Denys says that the procedure was successful and had the desired effect. However, although the patient had been given two successful blood transfusions over the course of several months, he died after he received a third transfusion. Denys reports that only the first two were endorsed by him and his colleagues and that the third was only performed "at the instant request of his Wife" (1668) after the patient's symptoms returned despite the initial success. In his letter, Denys blames the patient's own behaviour for the symptoms reoccurring:

(11) You already know, that the Transfusion of Calves-bloud did so temper the excessive heat of the Bloud of the Mad man, who for four moneths had run naked up and down the streets night and day; that he fell asleep two hours after the Operation, and that after ten hours sleep he awaken'd being in good sense, and that he remain'd in that good condition about two Moneths, until **the too frequent company with his wife, and his debauches in Wine, Tobacco, and Strong Waters** had cast him into a very violent and dangerous Feaver.

(Denys, PT 3(36), 1668: 710–5)

Later in the letter, Denys also casts doubt over the patient's wife, who he accuses of insisting on a third blood transfusion because she was having extramarital affairs and wished to harm her husband.

The reason why the patient and his wife are attacked particularly vehemently is likely to be because the patient died after an untested, new operation, rather than from more established medical procedures, and therefore the stakes are higher for Denys and his colleagues who set out to prove that blood transfusions could be safely performed on humans. By providing all of these details about the patient's life, Denys aims to show that blood transfusions are safe despite this one particular patient's death. Consequently, in addition to protecting his own reputation, Denys is also defending the reputation of blood transfusions. The first blood transfusions on a human leading to a death was undeniably the worst possible outcome for Denys and other proponents of the operation. However, by sharing this narrative with the Royal Society, Denys had the chance to manage how the events are perceived in England. Therefore, he seems to present their experiment in a more favourable light by trying to prove that the patient caused his own death, and that the skills of the people performing the transfusion were not at fault.

Overall, although treatment narratives are used to give the reader detailed instructions for diagnosing and treating illnesses, they also showcase the writer's medical knowledge. Sharing the range of different cases they have come across over the course of their careers adds to the trustworthiness of their medical advice and makes their text more authoritative. In treatment narratives, the writer's use self-mentions to present themselves as skilled practitioners who have the expertise to either effectively treat complicated cases or to give an accurate prognosis.

4.   Narratives of discovery

Whereas treatment narratives emphasise that the medical information shared by the author is reliable and often well-established, narratives of discovery focus on the process of research and the methods the author used to gather information. They are primarily concerned with gaining more information about rare medical conditions and new, untested remedies and procedures. In addition to giving the reader the author's findings, the self-mentions in these narratives aim to present the author as a reliable observer who is able to gather trustworthy evidence.

This is exemplified by the way William Harvey (1578–1657) describes his breakthrough discovery on the circulation of blood which went against the dominant Galenic theory of circulation (French 2004). In the English translation of *Anatomical Exercises* (1653), Harvey explains in a chapter entitled "The Causes which mov'd the Author to write" that the process he went through to get to these findings was painstaking:

(12) When **first I applyed my mind, to observation**, from the many dissections of Living Creatures as they came to hand […]; **I straightwayes found it a thing hard to be attained**, and full of difficultie, so with Fracastorius **I did almost beleeve, that the motion of the Heart was known to God alone** […]. At last using daily more search and diligence, by often looking into many and severall sorts of creatures, **I did beleeve I had hit the nail on the head**, unwinded and freed my self from this Labyrinth, and **thought I had gain'd both the motion and use of the heart**, together with that of the arteries, which I did so much desire […]. (Harvey, *Anatomical Exercises*, 1653)

As can be seen in the above passage, Harvey's narrative focuses on the internal, mental journey of acquiring new knowledge. In addition to impressing upon the reader the difficulties encountered, the narrative also emphasises that the conclusions Harvey's text presents are based on several observations gained through "daily more search and diligence". As Harvey's text is introducing a new theory to his readers, demonstrating that it is based on appropriately employed empirical methods adds to the reliability of his claims. Harvey concludes his narrative by saying that he decided to write his treatise after giving a lecture on his observations "which, as it commonly falls out, pleased some, and displeased others" (1653). Therefore, explicitly explaining that his findings come from extensive research are the result of a long, difficult process show the reader that Harvey is not making these controversial claims flippantly. Rather, Harvey demonstrates that he has followed the appropriate steps of empirical research by verifying his theory through multiple experiments, which adds to the authoritativeness of Harvey's conclusions.

Findings that are new or unexpected are presented in a similar manner in other contexts as well. For example, in 1686, Edmond King, writing to the Royal Society, reports the results of dissecting the body of "Mr. Robert Bacon, Master of Arts, of Corpus Christi Colledge in Oxford". King first explains that Bacon's family asked him to perform the post-mortem because they wanted to understand the cause of Bacon's strange and erratic behaviour in the years leading up to his death, which included moving furniture around from room to room as well as a "Canine" appetite and a desire to "have his Head rubb'd many times in the day". King records the findings of the dissection in great detail, covering the condition of the entire body. When he reaches Bacon's brain, King discovers an unexpected mass:

(13) The Glandula Pinealis firm and fair, well colour'd to look on, of the exact Figure, and ordinary sise: Feeling of it, and **finding it harder than ordinary** […] I prest it, and found in it a Stone in a film, or rather a petrified Gland in a film; I took out the Stone, and kept it as a great Raritie; I do not remember I ever heard of such a thing before, I am sure of all the Brains I have dissected (**and I may say I have dissected more than an hundred**) I never saw such a one.

(King, PT 16(185), 1686: 228–231)

King's narrative describes the methodical process of making this discovery and justifies that the finding is unusual by comparing it to the numerous other dissections he has performed. Following this, King goes on to detail the damage he observes in the surrounding area of the brain. The clear implication of King's findings is that the stone found in the brain was the cause of Bacon's behaviour. However, King himself does not offer this conclusion but rather finishes his report by emphasising his role as an impartial observer:

> (14) Thus having told you **Matter of Fact**, attested by his own Relations, (who were Eye Witnesses of it) **I leave these my Observations upon the Dissection of his Body to the Consideration of the more Curious and Inquisitive**.
>
> (King, PT 16(185), 1686: 228–231)

Overall, King focuses on showing that he uses the proper methods of observation to gain this information about Bacon's condition and that his part in this is purely to record what he observes rather than to prove or disprove any hypothesis about its underlying causes. The emphasis on his impartiality and his extensive experience of dissecting human brains present his findings as reliable and unbiased.

These narratives from Harvey and King are markedly different from the treatment narratives discussed above, as Harvey's text focuses on anatomy in general and King's on dissecting a dead body. However, within the sample examined in this case study, there are two examples where the events described in the narrative are related to treating a patient but instead of the outcome of the treatment, the writers are primarily concerned with gathering information about the patient's condition. Both of these narratives come from the *Philosophical Transactions*. In 1668, the journal published an account of Henry Stubbes's (1632–1676) journey to Jamaica in 1662 during which he held the office of king's physician (Feingold 2008). In this account, Stubbes shares multiple observations about the journey and the patients he treated. In the narrative passage studied here, Stubbes talks about two sailors who suffered from calenture, or heat stroke, and notes that the amount of information he can share about the disease is limited because the patients recovered quickly:

> (15) In our ship, two had the disease, so much talk'd of, called the Calenture; **concerning the progress of which disease I can say nothing; for they were thus cured presently**. I was talking with one of them, and on a sudden he beheld green leaves, as he imagin'd, floating on the sea, which yet was Azure-coloured: after that, he began to admire the fine woods, which he fancied to be near us. I immediately gave him a Vomit of the Glass of Antimony in Sack; which no sooner had wrought its effect, but all those imaginations vanished.
>
> (Stubbes, PT 3(36), 1668: 699–709)

The primary focus of this short narrative is describing the symptoms of the disease, whereas the treatment itself is only mentioned briefly. This makes the passage fit

better under discovery narratives than treatment narratives, even though the actual events of the narrative describe how one of the patients was cured. Although this passage does illustrate Stubbes's ability to effectively treat the sailors, the main purpose to share information about how calenture affected them.

Whereas in Stubbes's narrative the patients are treated successfully, in Daniel Turner's (1667–1741) report published in 1694, Turner discusses a failed attempt to treat a 3-year-old child suffering from rabies. Turner first describes the circumstances under which the child was bitten by a dog and then describes the progress of the disease and the dog-like behaviours the child exhibited before his death:

(16) I took the occasion (**out of Curiosity**) to present a Looking-glass before him, but found him so extreamly disturbed thereat, that I immediately took it away: He was no sooner sensible of the Reflection, than that he threw his Head backwards with great violence, and continued barking, and snapping at every thing near him: In the Evening […] he sunk under the Oppression of these cruel Symptoms. **I would very desirously have opened his Body, but it was forbidden by his Parents**. The Abdomen, I perceived, was excessively inflated, his Limbs convuls'd, and the superfice of the Body of a livid colour[…].

(Turner, PT 18(207), 1694: 24–25)

Compared to the other narratives of failed treatments discussed above, Turner's shows a lack of concern for the outcome, and the focus is solely on the information that can be gathered from the case. Turner does not attempt to emphasise accurate prognosis or offer any other explanation that could protect his reputation. Instead, Turner simply remarks on the fact that he was not able to perform a post-mortem due to the parents' objections, and then goes on to list the limited observations he was able to make without dissecting the body.

It is necessary to note here that these narratives from Stubbes and Turner were written for the *Philosophical Transactions*, and therefore the focus on sharing new information over anything else is to be expected, as that was the explicit mission of the journal (Hiltunen 2010). Nevertheless, all the four narratives discussed here illustrate an authoritative writing style that is different from the treatment narratives. The self-mentions still position the writer as an authority, but their authority comes from being able to accurately observe and record new findings. Whereas treatment narratives demonstrate the professional experience and medical knowledge the author already has, the narratives of discovery present the writer as an authority because they are familiar with the appropriate methods of observation and therefore capable of gathering new information reliably. In these narratives, the medical professional takes on the role of an impartial observer rather than a practitioner who comes into the situation already knowing everything they need to know in order to understand the patient's condition.

5. Discussion

The 31 narratives examined in this case study suggest that self-mentions in medical narratives are generally used to present the medical writer as an authority, but the writers take on different kinds of authoritative roles. Two different types of professional identity emerge from these narratives: the doctor as a practitioner and the doctor as a researcher. These roles are not mutually exclusive and could be, and mostly were, occupied by the same person, as can be seen in Examples (15) and (16) above where the doctors are simultaneously treating a patient and gathering new information. Rather, the difference lies in how the medical writers choose to present the events to the reader and what kind of information they focus on. In the case of treatment narratives, the writer is presented as a highly skilled practitioner who already possesses the necessary medical knowledge to diagnose and treat patients as well as predict the outcome of the treatment. The aim of self-mentions in these narratives is to show that the writer's extensive professional experience makes their advice trustworthy, which overall makes their text an authoritative source of established medical knowledge. In the narratives of discovery, on the other hand, the writer is a scientist whose role is to gather reliable evidence. In these cases, the focus is not on what the writers already know but rather their ability to make accurate observations, which encourages the reader to trust the validity of their findings. Both kinds of narrative present the author's experience as evidence of their skill, but the way their skills manifest within the narrative is different.

All in all, this sample includes 18 treatment narratives and 10 narratives of discovery, as well as three narratives that do not fit either category.[2] While the narratives of discovery and their focus on experimental methods are particularly linked to the empirical worldview, the explicit presence of the author's own experiences in the text is overall evidence of valuing professional experience over the teachings of past authorities. These self-mentions can, however, also coexist with elements of scholastic reasoning, such as Bullein's reference to Hippocrates (cf. Example 8 above). Nevertheless, whether the writer's authority is based on applying their medical knowledge and education or their observational skills, ultimately the focus is on using the author's experiences to prove that something is effective or true, as opposed to using their knowledge of past authorities to support their claims.

---

2. Two of these narrative passages come from John Caius's *Against Sweatyng Sicknesse* (1552), and both of these are autobiographical descriptions of Caius's personal background and education. The third comes from Thomas Gale's *Institution of a Chirurgian* (1563), which uses a fictional narrative of three surgeons meeting each other in a park as a framing device for his treatise on surgery.

The kind of authoritative positions the writers assume seems to also vary based on the purpose and the intended audience of the text. For instance, 50% of the treatment narratives come from surgical treatises which aim to give the reader practical guidance on how to perform certain procedures. By contrast, 6 out of the 10 discovery narratives come from the *Philosophical Transactions*, which focuses on disseminating new findings and scientific breakthroughs. Although further study would be required to make any reliable generalisations, such as collecting a larger sample by performing corpus searches on all 1st person singular pronouns that collocate with past tense verbs, the results of this study suggest that the audience and aims of the text influence how the medical writers construct their professional identity.

6. Conclusion

As has been discussed above, qualitative analysis of corpus linguistic data makes it possible to systematically study the pragmatics of historical texts. While corpus linguistic methods provide insights into overall trends, examining the issues highlighted by the quantitative results within their textual context can expand on the corpus findings and clarify the underlying pragmatic functions of the author's language use. In the case of the present study, the 31 individual narratives examined here provide a picture of how medical writers communicate their authority to their readers through self-mentions in narrative passages.

Although the writers present themselves as authoritative medical professionals in all of the narratives, this case study has demonstrated that they take on different roles based on the focus of their text. Whereas in treatment narratives the writers highlight the medical knowledge they already possess, in narratives of discovery they present themselves as authoritative because they know how to perform experiments and collect information through careful observation. In other words, the writers' self-representations range from an all-knowing medical practitioner to a researcher whose expertise lies in being capable of gathering new scientific information. Nevertheless, both types of narratives represent an empirical worldview, as they use the author's professional experiences and observations as evidence rather than relying on the scholastic method of inquiry. Therefore, using the narrative discourse mode to share their past experiences with their readers appears to be one of the ways in which early modern medical writers navigated the changes in scientific discourse, as it allowed them to establish themselves as an authority in line with the newly dominant empirical worldview.

References

Electronic resources

Early Modern English Medical Texts (EMEMT). Taavitsainen, Irma, Päivi Pahta, Turo Hiltunen, Martti Mäkinen, Ville Marttila, Maura Ratia, Carla Suhr, and Jukka Tyrkkö. 2010. In *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen and Päivi Pahta Amsterdam/Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/z.160

Primary sources

Bullein, William. 1562. *Bulleins Bulwarke*. Originally London: Ihon Kyngston. [STC (2nd ed.) 4033]. Source copy: Facsimile edition, *The English experience No. 350*, Amsterdam and New York: Theatrum Orbis Terrarum and Da Capo Press, 1971.

Clowes, William. 1596. *Booke of Obseruations*. London: Edm. Bollifant for Thomas Dawson. [STC (2nd ed.) 5445.5]. Source copy: BL 7481.d.5.

Clowes, William. 1602. *The Artificiall Cure of that Malady called in Latin Struma*. Originally London: Edward Allde. [STC (2nd ed.) 5446]. Source copy: Facsimile edition, *The English experience No. 238*, Amsterdam and New York: Theatrum Orbis Terrarum and Da Capo Press, 1970.

Colbatch, John. 1698. *Novum Lumen Chirurgicum*. London: printed by J.D. for D. Brown. [Wing (CD-ROM, 1996) C4999]. Source copy: BL 1608/5817.

Couch, Robert. 1680. *Praxis Catholica*. London: printed for Robert Harford. [Wing (2nd ed.) C6510]. Source copy: BL 1039.a.36.

Gale, Thomas. 1563. *Institution of a Chirurgian*. London: Rouland Hall [for Thomas Gale]. [STC (2nd ed.) 11529]. Source copy: WL 2504/B/1.

Harvey, William. 1653. *Anatomical Exercises*. London: printed by Francis Leach, for Richard Lowndes. [Wing (2nd ed.) H1083]. Source copy: BL E.1477. (2,3.).

Paré, Ambroise. 1634. *Workes of Ambrose Parey*. London: printed by Th: Cotes and R. Young. [STC (2nd ed.) 19189]. Source copy: BL C.112.h.9.

Philosophical Transactions. 1667–1668. Volume 2. Originally London: printed by T.N. for John Martyn. Source copy: Facsimile edition, *Philosophical Transactions, Volume 2, 1667*, New York: Johnson Reprint Corporation and Kraus Reprint Corporation, 1963.

Philosophical Transactions. 1668. Volume 3. Originally London: printed by T.N. for John Martyn. Source copy: Facsimile edition, *Philosophical Transactions, Volume 3, 1668*, New York: Johnson Reprint Corporation and Kraus Reprint Corporation, 1963.

Philosophical Transactions. 1686–1692. Volume 16. *Originally* London: printed by Joseph Streater and to be sold by Sam Smith. Source copy: Facsimile edition, *Philosophical Transactions*, Volume *16*, 1686–1692, New York: Johnson Reprint Corporation and Kraus Reprint Corporation, 1963.

Philosophical Transactions. 1694. Volume 18. *Originally* London: printed for S. Smith and B. Walford. Source copy: Facsimile edition, *Philosophical Transactions, Volume 18, 1694*, New York: Johnson Reprint Corporation and Kraus Reprint Corporation, 1963.

Secondary sources

Connors, Robert J. 1981. "The Rise and Fall of the Modes of Discourse." *College Composition and Communication* 32 (4): 444–455. https://doi.org/10.2307/356607

Cook, Harold J. 2008. "Colbatch, Sir John (bap. 1666?, d. 1729), Physician." In *Oxford Dictionary of National Biography*. https://www-oxforddnb-com.ezproxy.lib.gla.ac.uk/view/10.1093/ref:odnb/9780198614128.001.0001/odnb-9780198614128-e-5833

Feingold, Mordechai. 2008. "Stubbe [Stubbes, Stubbs], Henry (1632–1676), Author and Physician." In *Oxford Dictionary of National Biography*. https://www.oxforddnb.com/view/10.1093/ref:odnb/9780198614128.001.0001/odnb-9780198614128-e-26734

French, Roger. 2003. *Medicine before Science: The Business of Medicine from the Middle Ages to the Enlightenment*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511614989

French, Roger. 2004. "Harvey, William (1578–1657), Physician and Discoverer of the Circulation of the Blood." In *Oxford Dictionary of National Biography*. https://www.oxforddnb.com/view/10.1093/ref:odnb/9780198614128.001.0001/odnb-9780198614128-e-12531

Gotti, Maurizio. 2011. "The Development of Specialized Discourse in the *Philosophical Transactions*." In *Medical Writing in Early Modern English*, ed. by Irma Taavitsainen, and Päivi Pahta, 204–220. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511921193.012

Hiltunen, Turo. 2010. "Category 6. *Philosophical Transactions*." In *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen, and Päivi Pahta, 127–132. Amsterdam/Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/z.160.04cat6

Hyland, Ken, and Polly Tse. 2004. "Metadiscourse in Academic Writing: A Reappraisal." *Applied Linguistics* 25 (2): 156–177. https://doi.org/10.1093/applin/25.2.156

Jajdelska, Elspeth. 2018. "The Medical Paratext as a Voice in the Patient's Chamber: Speech and Print in *Physick for the Poor* (1657)." In *Medical Paratexts from Medieval to Modern: Dissecting the Page*, ed. by Hannah Tweed, and Diane Scott, 111–122. Cham, Switzerland: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-73426-2_7

Jucker, Andreas H. 1995. *Historical Pragmatics: Pragmatic Developments in the History of English*. Amsterdam/Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/pbns.35

Myhre, B. A. 1990. "The First Recorded Blood Transfusions: 1656 to 1668." *Transfusion* 30 (4): 358–62. https://doi.org/10.1046/j.1537-2995.1990.30490273445.x

Ollikainen, Karoliina. 2019. Authority in Early English Medical Writing (1375–1700): A Corpus Linguistic Study. PhD dissertation. University of Glasgow.

Pahta, Päivi, and Irma Taavitsainen. 1998. "Vernacularisation of Medical Writing in English: A Corpus-Based Study of Scholasticism." *Early Science and Medicine* 3 (2): 157–185. https://doi.org/10.1163/157338298X00266

Pomata, Gianna. 2010. "Sharing Cases: The *Observationes* in Early Modern Medicine." *Early Science and Medicine* 15: 193–236. https://doi.org/10.1163/157338210X493932

Siraisi, Nancy G. 1990. *Medieval and Early Renaissance Medicine: An Introduction to Knowledge and Practice*. Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226761312.001.0001

Smith, Carlota S. 2003. *Modes of Discourse: The Local Structure of Texts*. Cambridge/New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511615108

Taavitsainen, Irma. 2002. "Historical Discourse Analysis: Scientific Language and Changing Thought-Styles." In *Sounds, Words, Texts and Change. Selected papers from 11 ICEHL, Santiago de Compostela, 7–11 September 2000*, ed. by. Teresa Fanego, Belén Méndez-Naya, and Elena Seoane, 201–226. Amsterdam/Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/cilt.224.14taa

Taavitsainen, Irma. 2018. "Historical Corpus Pragmatics." In *Methods in Pragmatics*, ed. by Andreas H. Jucker, Klaus P. Schneider, and Wolfram Bublitz, 527–554. Berlin/Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110424928-021

Taavitsainen, Irma, and Susan M. Fitzmaurice. 2007. "Historical Pragmatics: What Is It and How We Do It" In *Methods in Historical Pragmatics*, ed. by Susan Fitzmaurice and Irma Taavitsainen. Berlin/New York: Mouton de Gruyter. https://doi.org/10.1515/9783110197822.11

Turner, Marion. 2016. "Illness Narratives in the Later Middle Ages: Arderne, Chaucer, and Hoccleve." *Journal of Medieval and Early Modern Studies* 46 (1): 61–87. https://doi.org/10.1215/10829636-3343123

Wear, Andrew. 2000. *Knowledge and Practice in English Medicine, 1550–1680*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511612763

Werlich, Egon. 1976. *A Text Grammar of English*. Heidelberg: Quelle und Meyer.

How old is old?

The discourse of "good" ageing in nineteenth-century self-help medical texts

Kim Grego

University of Milan

Life expectancy in developed countries has increased significantly over the past two centuries. This study contributes to the existing literature on how "old age" was defined and described at the onset of this development, namely in nineteenth-century British self-help texts, directed at advising people on medical practices to age well, or cure specific ailments. Eighteen texts were retrieved from the online library of the Wellcome Collection (https://wellcomecollection.org/), employing the search term "old age". A selection of these were analysed in depth, using specialised discourse analysis and critical genre analysis and also relying on socio-historical insights. These texts represent a diverse range of genres and contents employing various linguistic strategies and, thus, contributing to increasing the hybridity of the self-help manual genre in a time of great scientific and social change.

Keywords: elderly, self-help manuals, nineteenth century, English, Britain, medical discourse, genre analysis

1. Introduction

Just as life expectancy in developed countries has increased significantly over the past two centuries, so has the concept of "old age" undergone changes as regards both its chronological definition and the description of what it entails and the expectations that surround it. Early modern medicine in England has been widely investigated, and so has the eighteenth century as a historical period (e.g. Porter 1995, 2002, Bynum 1990, Kramer et al. 2003, Ottaway 2004, Troyansky 2005, Yallop 2013, Taavitsainen et al. 2014, and Taavitsainen and Hiltunen eds. 2019). The eighteenth century, in Ottaway's (2004: 6) words,

> was a time that has been traditionally associated with the birth of both the demographic transition and the industrial revolution. As mortality and age at marriage dropped, fertility rose, and the population of England grew both larger and younger by the end of the century.

By the nineteenth century, an increase in the average life span was a well-established reality.

Porter observes that in the eighteenth century, for the first time, "alongside the personal and scholarly writings of practitioners, there is no lack of medical diaries, letters and auto-diagnoses penned by sufferers themselves" (2002: 289). Joan Lane, analysing "[t]he diaries and correspondence of patients in eighteenth century England", reports that the diarists she researched "unless hypochondriacs tended to need medical attention more at certain periods of their lives, especially old age (2002: 213). Old age is often mentioned in medical documents, whether in references to doctors themselves, e.g. "Dr Hale […] 'an old physician of great repute' in 1748" (Ibid.: 233), or to patients, who were often classified in eighteenth-century self-help writings according to simple, traditional categories, "such as male and female, purity and impurity, infancy and old age" (Smith 2002: 256). In particular, this

> popular advice in practice was also broken down into advice for the old, the young, females […] (considered 'weak' or vulnerable, as were the young and old).
>
> (Ibid. 2002: 257)

Such vulnerability of the elderly, not only concerning their health but also reflected in their social condition, is precisely one of the reasons that make them interesting from a critical discourse studies viewpoint, focusing on discrimination in society.

Yet, old age vulnerability cannot be fully understood without looking back at the period when this weaker social category began to emerge as a subject of interest for medical writers as well as for publishers. As Porter (2002) well underlines,

> [t]he real watersheds are generally seen as occurring before or after the eighteenth century, with for example the Scientific Revolution, or the emergence of the family doctor in the early nineteenth century, or the coming of medical registration in 1858.
>
> (Porter 2002: 285)

It is precisely the latter period, from its beginning to the late nineteenth century, that the present investigation concentrates on.

The aim of this chapter is to study how "old age" was defined and described in nineteenth-century British texts, i.e. writings directed at laypeople by specialists, advising them on medical practices aimed at aging well, or at curing particular ailments in old age. The specific research questions pursued in this study are:

1. What was the contemporary understanding of *good* ageing?
2. What conditions were associated with it?
3. What was the discursive rendition of this stage of life?

2. Data and sources

For the purposes defined above, the Wellcome Collection was selected as a suitable repository of texts to examine. First, the online catalogue was searched for the collocation "old age" in English-language texts published in Britain[1] between 1800 and 1899.

   This search returned an initial number of 130 titles. After removing doublets, different editions of the same text, works published outside Britain, and English translations of foreign works, the remaining texts were manually screened to identify those dealing with medicine or health, and addressed to a lay audience rather than to specialists. This process resulted in the selection of 18 works, which were divided into loosely defined categories based on their addressees: (1) the aged, (2) young people or adults wishing to age well and (3) everyone (the perspective of the texts in this category is essentially philosophical). The categorisation is shown in Table 1.[2]

Table 1. Medical and health popular texts from the Wellcome Collection (1800–1899)

No.	Category	Subcategory	Text
1	Addressing the aged (1)		**Carlisle, Anthony (1817)**
2			Van Oven, Bernard (1853)
3	From infancy to old age,	*General* (2a)	Alcott, William Andrus (1830)
4	through one's prime		Medicus [pseudonym] (1830)
5			**Anonymous (1832)**
6			Jones, William (1858)
7			Whitmarsh, William Michael (1870)
8		*Specific* (2b)	Murphy, Joseph (1811)
9			Tweed, John (1820)
10			West, Francis (1827)
11			Imrie, William (1835)
12			**Black, James (1840)**
13			Canton, Alfred (1851)
14			Browning, John Dudley (1884)
15			Ballin, Ada S. (1893?)
16	Philosophical (3)		Sinclair, John (1807)
17			Bernard, Thomas (1813)
18			**Forster, Thomas (1829)**

---

1. Most were published in England. Three of the four manuals analysed in depth were published in London (see Table 1), but their authors came from different parts of the country; e.g. Dr Black, was a Glaswegian who practiced in Scotland. This discrepancy between the places of publication (the most prestigious houses would be in London) and the authors' places of origin and/or practice, together with the national level of circulation of these manuals, suggests "Britain" might well be used as a geographical indication.

2. Full titles are found in the Primary Sources in the References.

The analysis will focus in particular on one representative text per (sub)category (cf. titles in bold in Table 1). They are discussed in the chronological order: Carlisle (1817),[3] Anonymous (1832),[4] Black (1840)[5] and Forster (1829).[6]

3.   Methods

This study adopts the perspectives of specialised discourse analysis and critical genre analysis, and relies on socio-historical insights. The eighteen texts were manually searched for definitions of what age their authors meant by "old". In particular, the four sample texts were studied in detail and analysed in relation to the period developments in the English language in general (Görlach 1999; Kytö et al. 2006), paying special attention to the history medical texts (Porter 2002, 1995) and their social context (Lawrence 1994, Lane 2001, Waddington 2011, Worboys 2011).

The English language of the 1800s, although not dramatically different from that spoken and written in the previous century, shows a number of peculiarities described, for instance, by Görlach (1999). The most interesting developments took place in lexis and genres. The nineteenth century is characterised by genres becoming increasingly diverse, especially in scientific discourse, making it a period of interest both from synchronic and diachronic perspectives (Kytö et al. 2006: 4). The key concepts emerging from this study are the rise in the production and publication of scientific texts, the development of related genres, the growing popularisation of science, and the relevance of these changes to the synchronic level.

Socio-historical studies of the role of medicine in society were also consulted, especially as regards to changes in Britain where "in the two centuries from 1750 to 1950 medical practice itself changed out of all recognition" (Lane 2001: vii).[7] Lawrence (1994) remarks that during the eighteenth century, "most hospitals did not admit the chronically (long-term) sick, the aged, the mad or children" (Ibid.: 16); they were treated at home and not necessarily by qualified physicians. It was not until the nineteenth century that this "motley crew" of socially weak individuals were sheltered in workhouses in the wake of the Poor Law Amendment Act of 1834 (Ibid.: 32).

---

3.   *An Essay on the Disorders of Old Age.*

4.   *Twenty Minutes' Advice.*

5.   *A Non-Medical Treatise. A Manual on the Bowels.*

6.   *Medicina Simplex.*

7.   Cf. Waddington's (2011: 2) suggestion of the French Revolution representing "a critical turning point in the history of medicine".

This study proposes a discourse-based interpretation with a point of departure in Foucault's (1963/2003) "medical gaze", through contemporary views of language in specialised settings (Gotti et al. 2015), and Critical Genre Analysis (CGA) (Bhatia 2017). Foucault's influence is easy to justify, as his gaze lingered long on the sick and on the development of medical practice, reviewing the history of the clinic as a place and as a discursive entity (1963/2003). His work has deeply influenced discourse studies, especially those focusing on the medical domain (Foucault 1980; Jones and Porter 2001). The papers collected in Gotti et al. (2015) provide an ample review of studies on medical communication. Maci et al. (2015: 9) observe how "specialists employ as many registers as possible according to the many different options available so as to target the various audiences they need to address" to achieve "a more 'open genre network'", and these considerations can be applied to the historical context of nineteenth-century popular medical texts. Similarly, Bhatia's CGA can provide tools for demystifying professional practice through the analysis of genres (2017: 9), and as such can shed new light on nineteenth-century self-help manuals for good ageing.

4. Analysis

4.1 Comparing the four sample texts

The four works selected for closer scrutiny are genuine self-help texts for the lay-people, as openly stated in their prefaces or introductions. Their authors came from heterogeneous backgrounds and show diverse features (see Table 2).

Table 2. The four sample texts

Text	Length (pages)	Price	Publisher	Author's background	Audience	Topic	Genre(s)
Carlisle, Anthony (1817)	120	5s	Longman, London	M.D.	Old people	Seniors' health	Health manual
Forster, Thomas (1829)	296	3s	Haddon and Fenton, Colchester	Non-practising M.D.	Adults	General health	Health manual but esp. moral guide, catechism calendar, religious treatise
Anonymous (1832)	39	1s	William Kidd, London	Expert? Editor?	Adults	General health	Practical hands-on short manual
Black, James (1840)	248	5s 6d	Longman, London	M.D.	From infants to the aged	Bowel health	Health manual with specialist's observations and comments

4.1.1 Pages, prices, publishers

These texts reflect the wide variety of self-help manuals in the nineteenth century. Differences are found in the numbers of pages ("Length" in Table 2) and the prices, with Anonymous (1832) being the only one displaying its price on the front cover (the others were retrieved from book reviews). Both Carlisle (1817) and Black (1840) were published by Longman, the prestigious London house, active since 1724.[8] The assumption is that the two manuals published by Longman enjoyed more scientific prestige than the others. Forster (1832)[9] was printed by Meggy and Chalk, High Street, Chelmsford, and "sold by Keating and Brown, and Booker, London; and by all booksellers in town and country" (the frontispiece). The 1832 edition was printed by "Haddon and Fenton, Colchester, sold in London by Keating and Brown, in Colchester by Haddon and Fenton and in Chelmsford's High Street by J. Marsden". Although these smaller publishers did not enjoy the same prestige as Longman did, they were quite active with a tight sales network from London throughout all Essex. Finally, Anonymous (1832) was published by the scarcely known W. Kidd in London.

The author of the earliest manual, Anthony Carlisle (1768–1840), was a renowned English physicist and a surgeon, a Councillor and then President of the College of Surgeons, a Fellow of the Royal Society and Surgeon Extraordinary to King George IV (Cole 1952). He seems to have enjoyed visibility in the high circles, and maybe among the lower classes, too.

Thomas Ignatius Maria Forster (1789–1860) was better known as a physicist pre-eminently interested in astronomy and a Fellow of the Royal Astronomical Society. Medicine was just one of his many interests, and the subject he had graduated in at Cambridge, but "he never practised" (ODNB 2004; Janet Browne, s.v. 'Forster, Thomas Ignatius Maria'). Incidentally, he was also a Catholic convert, a phrenologist and a vegetarian *ante litteram* (Ruston 2005). His peculiar background and views on life heavily influenced the treatise considered in this study, as shown in later sections.

The third source, the one-shilling, 39-page booklet *Twenty minutes' advice*, is anonymous. However, by looking at its publisher, W. Kidd, some speculations are pertinent. The *Oxford Dictionary of National Biography* (2004; M. G. Watkins, revised by H. C. G. Matthew, s.v. 'Kidd, William') lists a William Kidd (1803–1867) from London, who joined the publishing profession as a young apprentice, but was always particularly interested in natural sciences, and "entered business on his own account, and had shops successively in Chandos and Regent streets,

---

8. It was famous for the publication of Samuel Johnson's *Dictionary* (1755).

9. The first edition from 1829 was unavailable.

London" (Ibid.). This is precisely what appears on the frontispiece of the 1832 manual (see Figure 1). The publisher should then definitely be William Kidd from Chandos Street in London, but we also learn that "[w]hile at Chandos Street Kidd published a *Guide to Gravesend*, *Popular Little Secrets*, and other short essays written by himself" (Ibid.). This suggests that Kidd himself could be the author of the booklet from 1832. Although not a doctor of medicine, his background knowledge of the natural sciences seems solid enough for him to have produced a self-help text of this kind.

The author of the last book, James Black (1788/9–1867), was a well-known practicing physician, who lectured at the Royal School of Medicine and was a Fellow of The Royal Society of Edinburgh (ODNB 2004; C. W. Sutton, revised by Patrick Wallis, s.v. 'Black, James'). Doubtless the prestigious publisher matches the accomplished authors, and so does, in its own way, the Anonymous (1832) booklet with its supposed author and self-publisher. The less usual combination is the one connecting Thomas Forster, a well-versed graduate in medicine, and Meggy and Chalk / Haddon and Fenton, small-town, high-street printers-turned-publishers.
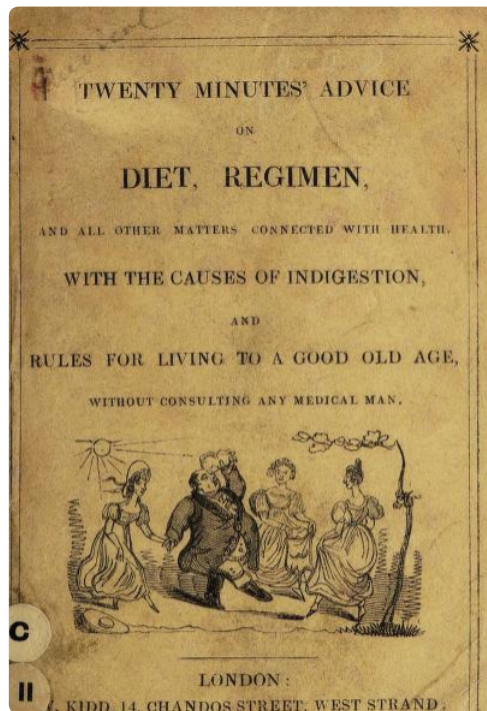
Figure 1. Frontispiece of Anonymous (1832) (Wellcome Collection, Public Domain Mark)

4.1.2   *Public and purpose*

Turning to the specific addressees and purposes, Carlisle (1817) begins with a dedication, not to the aged public whom he targets, but to his fellow professionals. The dedication is a formal (sub)genre in itself, employing well-established formulas ("Respected Colleagues", "all men of education", "the Healing Art", "laws of Nature", "Your much obliged, And faithful Servant"; Carlisle, 1817: iii–iv), and the ritual apologetic statement: "I submit this short exposition of a part of my professional sentiments to your competent and equitable judgement" (Ibid.: iii). His solid scientific background, rooted in the experimental method, emerges in his belief that medicine should be "established upon rational and scientific foundations [...] and constantly subject to the corrections of progressive Experience and accumulating Observations" (Ibid.: iv). In the incipit of the treatise itself, though, Carlisle lays out his intended aim and audience, declaring that it might be

> (1)   unnecessary to trouble the Public with any further additions [to previous medical tracts on longevity], were it not apparent, that many of those books are too voluminous for *common Readers*.   (Carlisle 1817: 1, emphasis added)

He then indicates that "[t]he present Essay is addressed principally to *persons already advanced in years*, and does *not* embrace the course of regimen which *from an early period* is best calculated *to secure a long life*" (Ibid.: 2, emphasis added), thus going so far as to specify for whom his book is *not*.

Forster's *Medicina Simplex* (1829) is subtitled *The pilgrims waybook, being an enquiry into the moral and physical conditions of a healthy life and happy old age: with household prescriptions*. Even without being aware of the author's curious life interests, the words *pilgrims*, *waybook* and *moral* do pertain to a semantic area close to things philosophical, religious or at least spiritual. Other words, like *physical*, *healthy* and *prescriptions*, point to the medical practice. This is a hybrid text that begins with an "Introductory analysis" that resembles a preface (Forster 1829: xi):

> (2)   I proposed to discuss all the means at present known, of producing a sound state of the animal machine, and a *happy longevity*; in short, of conferring mirthfulness and health on *youth*, strength and wisdom on *manhood*, on *old age* ease and tranquillity, and in the *end* a dissolution without pain.
>   (Forster 1829: xi, emphasis added)

The animal machine is an ancient concept made more popular by the growing body of studies on anatomy of the previous three centuries, as well as Newtonian medicine. Foster's intent mixes a scientific view with philosophical concepts such as happiness. As his addressees, Forster indicates practically everyone: from the young (*youth*), to those in their prime (*manhood*), to the old (*old age*). He even mentions death (*the end*) and the wish that it may be painless.

The anonymous *Twenty minutes' advice* (1832) is expectedly the most curious sample. Apart from the anonymity of its author, several details emerge from both the frontispiece and the preface. In the former, the writer mentions "health", while at the same time underlining that it is "A non-medical treatise". This clear-cut distinction between health and medicine is relevant not only philosophically and professionally, but also in the light of the different profile of this book's author, who obviously could not (or would not) claim to be a doctor of medicine. In fact, the subtitle of the text ends by wishing that people *may live to a good old age, without consulting any medical man*. This implies the supposition that the author was not a physician. A classic preface of the text shares typical genre features e.g. an apologetic tone in phrases such as *little work*, *simply*, *this little manual* and *[i]ts pretensions are humble* (Anonymous 1832: iii). The addressees are people at large: "This little work having been written simply with a view to benefit *the Public*" (Ibid., emphasis added). Science is as apologetically brushed off ("nor does it undertake to enter at large on the discussion of scientific matters", Ibid.), in favour of what could be termed sound common sense, to which the anonymous author almost proudly appeals: "it, nevertheless, professes to contain all that is *necessary to be known* on the subjects of which it treats" (Ibid., emphasis original). Incidentally, the absence of any hedging strategy – indeed, boasting to "contain all that is necessary" – contradicts the humble pretensions expressed just before. A reference to money is also interesting as it states that books on the same subjects exist, but "it is to be regretted that their high prices preclude their being generally useful" (Ibid.).[10] This suggests that the general public could not afford them and enjoy their benefits ("generally useful"), and that the price sets them *socially* apart from the common people. This could also imply that the texts were so specialised that they could not be understood by laypeople. A blatant marketing approach is similarly evident in the frontispiece by the same anonymous author.

(3) […] author of "Twenty minutes' advice on the gout;" "Twenty minutes' advice on head-ach," &c.; "Twenty minutes' advice on corns and bunnions," &c. &c.!

(Anonymous 1832)

The booklet's last two pages, as well as the back cover, advertise other works published by William Kidd of 14, Chandos Street, West Strand.

Finally, Black (1840) shows a traditional structure: a preface followed by two parts, the former dedicated to discussing topics theoretically and the latter reporting the author's own practical experience. The title indicates it is a manual about the bowels and treatment of bowel diseases from infancy to old age. The scope is thus quite broad, but in the preface the author acknowledges that "it is unnecessary to

---

10. This is indeed confirmed by the prices reported in Table 2.

offer any lengthened apology for its publication" (Black 1840: iii). Nevertheless the text abounds in apologetic expressions addressed both to his own community of practice, and the public, at whom the book is ultimately aimed: "[t]he subject, which he *has attempted* to treat in a *very condensed* manner", "this *little* work", "[it] will render it *not unsuitable* for the perusal and attention of the junior members of the profession", "[t]he Author *only desires* to add", "*authorities*, which *he has taken pleasure* currently *to acknowledge*." (Ibid.: iii–iv, emphasis added). However, this is not just Carlisle's "common Readers", Forster's "people of all ages", or the Anonymous author's "general Public", but "the *intelligent* reader, [to whom] it will convey that information on the structure and functions of the most important and outwardly related organs of his body, and those rules for his personal hygiene, as may be easily comprehended, and practically followed" (Ibid.: iii–iv, emphasis added).

4.1.3   *Topics and genre*

The topics discussed in the manuals are general in all the cases, except for Black (1840), which focuses on bowel issues. In terms of genres, all texts can be safely defined self-help health texts, with a few caveats. Carlisle's work, the earliest and the one by the most illustrious author, is written in the classic expository style, without divisions into parts or chapters. It is preceded by a preface. Forster's book distinguishes itself firstly because, although also expository, it is divided into sections. Secondly, it shares the features of a moral guidebook, a catechism and a religious treatise. This is not surprising, given the author's biography; the text is therefore hybrid in nature, with a strong religious/philosophical component.[11] The anonymous booklet from 1832 is possibly the best incarnation of the self-help health manual genre: a short, practical and hands-on instructive text with simple and specific instructions. Lastly, Black (1840) uses an expository style in the first part of his book, but adds numerous observations from his experience in the second section. This is possibly consistent with the style of medical writing of the times, at least judging from the other 14 texts in the corpus and was, anyway, justified by the writer's great experience and prestige.

4.2   Defining old age

Returning to the initial conundrum of how to *delimit* and *define* old age, a manual survey was made of the eighteen texts in the corpus. The different time limits and definitions of old age encountered are summarised in Table 3.

---

11.   Indeed, it is listed by Rodd (1837) in his *Catalogue of Books for MDCCCXXXVII. Containing Theology and Ecclesiastical History. Metaphysics, Moral Philosophy and Education.*

Table 3. How old is *old*?

Cat.	No.	Author, year	When is old age?
1	1.	**Carlisle, 1817**	**"The age of Sixty may, in general, be fixed upon as the commencement of Senility", p. 13**
	2.	Van Oven, 1853	70 is old, 80 is very old, p. 58
2a	3.	Alcott, 1830	"[S]eventy – at which time caducity begins", p. 215
	4.	Medicus, 1830	"[G]reen old age" is at least 50, p. 175
	5.	**Anon., 1832**	**Not specified**
	6.	Jones, 1858	60, quoting Carlisle (1817), p. 255
	7.	Whitmarsh, 1870	N/A*
2b	8.	Murphy, 1811	When teeth are lost, a period "various and uncertain", p. 90
	9.	Tweed, 1820	Current lifespan set at 70–80 years, p. 23
	10.	West, 1827	Reference to Dr Jameson: meridian of life is 28; according to others, 35; according to author, by 40 eyesight decreases, p. 11
	11.	Imrie, 1835	Reference to Hunter: decay of teeth begins after 50 years, p. 19
	12.	**Black, 1840**	**"[W]hen the meridian of life is passed, say from 45 to 50 years of age", p. 115**
	13.	Canton, 1851	Not specified
	14.	Browning, 1884	Old sight: 45 women, 47 men, p. 53
	15.	Ballin, 1893?	Completion of growth in women not definitely fixed before 23, p. 157
3	16.	Sinclair, 1807	Longevity depends on profession, esp. long lived is the commercial class, pp. 549 ff.
	17.	Bernard, 1813	Studious men and *literati* tend to live longer, p. 135
	18.	**Forster, 1829**	**Not specified**

* This was the only text that could not be consulted, since it is only available in paper format at the Wellcome Library, which was closed due to the 2020 COVID-19 pandemic.

The books in Category (1) are the most "optimistic", indicating 60 and even 70 as the commencement of senility, probably because the approach was holistic and the audience is supposed to be already aged, so a "positive" perspective was probably adopted. On the other hand, the most diverse data is found in texts in Category (2b), which focus specifically on a single organ or medical issue. The age at which old age begins is low for teeth health and even lower for the eyesight. As we approach the end of the century, some gender distinctions also appear: Browning (1884) sets a different inception for a decrease in men and women's vision, while Ballin (1893?), which incidentally is the only one authored by a woman and the latest text published, focuses on women. Finally, also unsurprisingly, the philosophical treatises in Category (3) do not provide answers about when one becomes old (Forster 1829), or they differentiate according to the person's profession or activity (Sinclair 1807 and Bernard 1813).

Of the four sample texts investigated in detail, Carlisle (1817: 13) is the only one specifying an age for "the commencement of senility": 60 years. As a man of science, he nonetheless adds the phrase "in general", which is both a hedge and an acknowledgment that medicine is no exact science. Black (1840: 113) also gives a figure, or rather a range: "when the meridian of life is passed, say from 45 to 50 years of age". No specific reason was found to account for the ten-year difference in the two doctors' estimates, not even the almost quarter of a century separating the two manuals, considering that (even if it is merely a statistical datum) in 1833 the average life expectancy in the UK was 40.9 years, while in 1842 it had only increased to 41.[12] No indications as to "how old is old" arrive from either Forster (1829) or Anonymous (1832). It is possible that both took it for granted that old age existed ontologically, irrespective of chronological boundaries – the former from a philosophical perspective and the latter from mere common sense.

The specific terms and phrases employed by each author are reported in Table 4.

Table 4. Old age: definitions (absolute frequencies)

Carlisle (1817)	Forster (1829)	Anonymous (1832)	Black (1840)
– *old persons*: 8	– *old man*: 2	– *old men*: 8	– *the aged*: 3
– *the aged*: 8	– *the aged*: 1	– *old age*: 2	– *aged people*: 2
– *aged persons*: 2	– *aged catholics*: 1		– *aged individual*: 1
– *persons already advanced in years*: 1	– *aged persons*: 1		– *advanced in years*: 1
– *those who are stricken in years*: 1	– *old persons*: 1		
	– *old age*: 13		– *old age*: 9
– *old age*: 5	– *advanced stages of life*: 1		– *advanced life*: 3
– *advanced age*: 2			– *advanced age*: 2
– *advanced life*: 2	– *maturer life*: 4		– *advanced period of life*: 1
– *senility*: 2	– *senility*: 1		
– *a very advanced age*: 1	– *senectitude*: 1		– *advanced time of life*: 1
– *our declining years*: 1			
– *age*: 1			– *advanced years*: 1

Carlisle (1817) uses the widest range of terms. Considering that his book is 120 pages long, this might also imply that he paid some attention to the style of his writing. Forster (1829) uses *old man* (but not, e.g., *old woman*). Anonymous (1832) employed the smallest number of collocates. In between is Black (1840), who seems

---

12. Our World in Data, Life expectancy 1543 to 2015, UK, https://ourworldindata.org/grapher/life-expectancy.

to prefer noun phrases pre-modified by the deverbal adjective *aged* and by the polite alternative *advanced*. His lexical expressions are usually plain but transparent (*advanced in years*, *advanced period of life*). Black's manual is also the latest of the four manuals, and a slight yet definite change in the choice of terminology can be perceived between his and Carlisle's usages. Forster (1829), too, employs a varied selection of expressions, some with a classic flavour (*senility*, *senectitude*, *mature life*), others constructed similarly to Black's (*advanced stages of life*). In sum, all refer to the elderly as *old men* and to their stage of life as *old age*, without any reticence in the use of an adjective which tends to be avoided in current times but was evidently widely accepted in the nineteenth century. The only slight change toward a possibly euphemistic alternative, *advanced*, is found in Black (1840).

4.3 The ailments of old age and their remedies

What is somewhat clearer is the kind of conditions that seniors suffered from in nineteenth-century Britain and, especially, the suggestions provided to treat or alleviate them. The conditions of old age are several, as seen in the complete list reported in Table 1 and, especially, in Category (2b). The specific ailments appearing in the four sample texts are listed in Table 5.

Table 5. The conditions of old age

Carlisle (1817)	Forster (1829)	Anonymous (1832)	Black (1840)
– the head	– **deafness**	– **stomach**	[All **bowel** related:]
– the stomach	– **mind**	– **bowels**	– slow elaboration and defecation
– **the bowels**			– loose retention of anus
– the blood vessels			– flatulent colic
– the liver			– itching of the anus
– apoplexy			– tape worms
– palsy			– ascarides
– indigestion			– thread worms
– obstructions			– torpor of the brain
– inflammations			– paralysis
– jaundice			– retardation
– dropsy			– constipation
– teeth			– relaxation and atony of rectum
– eyes			and sphincter
– **ears**			

Carlisle (1817), who specifically addresses seniors and their ailments, features quite a long and varied collection of sicknesses. While Black (1840) focuses exclusively on bowel conditions, he is very specific in mentioning the issues of the entire digestive tract. Probably for this reason, the disorders he mentions often have multi-word names (*loose retention of anus*, *relaxation and atony of rectum and sphincter*), which confirms his tendency to use longer, more specialised and, ultimately, more contemporary terms. The other two manuals are directed at middle-aged people who want to age well; thus, they are less specific and concentrate on giving advice on prevention rather than on treatments. Forster (1829: 208), true to his dichotomic view of the world, focuses essentially on deafness as a bodily issue and on the decay of the mind (*old age becomes a second childhood*). Apart from his understanding of the human as a machine animated by a soul, Forster testifies to the development of cognitive issues in those reaching old age. For demographic reasons, this was possibly not a frequent occurrence in his time, (see Footnote 9), but one that would come to grow exponentially to become the social problem it represents today. Interestingly, Anonymous (1832) only discusses stomach and bowel diseases, separately. This differentiation, which was not sustained by Black the gastroenterologist, comes naturally to the author of *Twenty minutes' advice*, who sees and treats the two organs distinctly. He also mentions headache, but again only as caused by indigestion. Accordingly, it is noticeable that the contents of the anonymous booklet begin with the mentioned stomach and bowel aches, then move on to discuss their causes, i.e. meals (size and timing) and, subsequently, the pros and cons of the most common foods and beverages. Thus, the one uniting element that all four manuals share are the *ailments of the gastrointestinal tract*, which emerge, at least quantitatively, as those mostly afflicting the elderly.

To prevent or treat these ailments, the four self-help texts provide several suggestions, recipes and advice, the nature of which, summarised in Table 6, alternates between the expected and the surprising.

A real geriatrist *ante litteram*,[13] Carlisle has a wide variety of remedies to offer, which match the extended range of ailments he mentions. Among these, we find "judicious management", "diet" and "preventive measures". Forster's (1829) spirituality neatly emerges in his exposition of how to obtain good old age: he recommends following "higher principles" and relying on "religious expectation of future happiness", observing "fasts and feasts", the "Catechism" and, only as a last resource, a "Change of Diet". Anonymous (1832) recommends "regular exercise" (cf. Carlisle's "moderate exercise"), proportion in the quantity and quality of food, and "regular

---

13. Geriatrics emerged as an institutionalised branch of medicine in the first decades of the twentieth century.

Table 6. What is good for the old age (with page numbers)

Carlisle (1817)	Forster (1829)	Anonymous (1832)	Black (1840)
– *judicious management* (3)	– *higher principles* (201)	– *proper attention to the digestive powers of the stomach* (9)	– *To avert, or at least procrastinate this physiological precocity of the stomach and bowels, […] youth […] should endeavour to carry youth as far as possible into old age, by using the wisdom of old age in their youth* (136)
– *wholesome regulation of **diet**, clothing, exercise, and air*	– *certainty in the **religious expectation of future happiness*** (201)	– *taking **regular exercise** in the air* (9)	
– *blood letting* (9)	– *sanatory rules of the Calendar, in the **fasts and feasts*** (211)	– ***proportion the quantity of food** […], quality […] and take it, at regular intervals of six or seven hours, thrice during the day* (16)	
– *prudent regiment and **preventive methods*** (10)			
– *choice of food and cookery* (29)	– *The **Catechism** […] forms the basis of that disposition of mind which in senility constitutes the source of happiness* (214)	– *Avoid walking, immediately before, or immediately after any meal* (31)	– *avoid tight stays and corsets* (216)
– *dining in the middle of the day* (48)		– *regulate your bowels* (34)	– *avoid hot and irritating clysters* (216)
– *warm clothing* (61)		– ***be regular and temperate** in your habits* (34)	
– *temperate atmosphere* (61)	– *Rules of Change of **Diet*** (260)		
– *good ventilation* (61)		– *refrain from the use of Spirituous Liquors* (34)	
– *moderate exercise* (62)			
– *regular bowel movements* (63)		– *retire to rest early* (34)	
– *avoid dangerous surgery* (94)		– *rise betimes in the morning* (34)	

and temperate" habits. Black (1840), advising his readers only on bowel problems, mentions avoiding "tight stays and corsets" or "hot and irritating clysters", but he invites the young, especially, to "endeavour to carry youth as far as possible into old age, by using the wisdom of old age in their youth" – a paradox that remains, to this day, as wise as it is impossible to put into practice.

5.   Discussion of findings

Close reading of the corpus data reveals several interesting features. Linguistically small yet detectable changes seem to occur in the four manuals. Carlisle (1817) uses language that is closer to the eighteenth-century style, with an abundance of Latinate words, despite the manual's informative purpose: "apoplexy, palsy, indigestions, obstructions, inflammations" (Carlisle 1817: 9). This is not surprising, given that, "[a]lthough many writers state that the 19th century had overcome the Latinate style of Johnson, complaints about contemporary excesses did not cease" (Görlach 1999: 116). Forster (1829) also shares a taste for and (ab)use of Latinate lexis. On the one hand, this is due to the philosophical/religious nature of his book and, on the other hand, "the entire anglophone world of the 19th century exhibits texts whose authors preferred Latinate diction at least for certain registers much more than would be usual today" (Görlach 1999: 107). The expectation about Anonymous (1832), given the premises regarding its publisher, possible author, genre and hybrid commercial purpose, was that it would include less vocabulary of classic origin, but classical vocabulary is still widely used, although often in conjunction with synonyms, not necessarily of Anglo-Saxon origin. For example, a Latinate, if a more popular, term is provided in exchange for a Greek one: "Dyspepsia, or Indigestion" (Anonymous 1832: v). In this dyspepsia section, the author undertakes a thorough linguistic analysis of the word, with a clear intent of disseminating knowledge:

> (4)   According to our best authority, (Dr. Paris,) *Indigestion* is nothing more than a literal English translation of the Greek compound Dyspepsia; and it is defined by him to be a *primary disease, in which one or more of the several processes by which food is converted into blood, are imperfectly or improperly performed, in consequence either of functional aberration, or organic lesion.*
>
> (Anonymous 1832: 10, emphasis original)

This includes a quotation from an authoritative source: Dr John Ayrton Paris (c. 1785–1856), whose exact work is nowhere cited in the anonymous booklet.[14] The anonymous author adds the apophatic comment: "But the definition we leave to the learned, while we come at once to the marrow of the subject" (Ibid.). This creates a shared tone with his lay reader, with whom he states he shares a preference for "coming to the marrow of the subject" – using a medical term in a metaphor that was intended to lower the register to meet his audience's level.

By contrast, Dr Black is one such learned doctor and writer that Anonymous (1832) half respects and half ignores, who makes ample recourse to terms of classical etymology:

---

14. The reference made by Anonymous (1832) is from Paris' *Treatise on Diet* (1826: 213–214).

(5) we shall first take notice of derangements and morbid alterations of the intestinal *secretions* and *excretions*; next, alterations and lesions of the *mucous membrane*; and, lastly, *disordered* and *anormal* affections of the *muscular coat*, along with some relative notice of those of the *peritoneal membrane*, and *disordered defæcation*. (Black 1840: 53, emphasis original)

Black (1840) is also interesting because of its larger use of specialised multi-word terms compared to the other texts (see Section 4.3). The kind of terms employed by Black (1840) – *advanced time of life*, *flatulent colic*, *itching of the anus*, *tape worms*, *thread worms* and *torpor of the brain* – seem to fall precisely into the blurry category of "lexical strings" or Görlach's (1999) "transparent compounds": not perceived and thus not recorded as such. The trend continues into contemporary scientific English word-formation, with the exception that some such terms would indeed make their way into specialised dictionaries.

According to Görlach (1999: 148), dedication forms in prefaces became less elaborate after 1840. Evidently, the four sample texts considered in this study still suffer from the residue of the earlier centuries' sycophancy, but that would be perfectly in line with Görlach's proposed timeline, since the last of the manuals, Black's, which sports one such formal preface, is precisely from 1840. Overall, the closer inspection of the sample texts highlighted both standard (Carlisle 1817, Black 1840) and hybrid (Forster 1829) or original (Anonymous 1832) realisations of the self-help treatise. This, again, matches Görlach's observation that

> the combination of a well-defined text type[15] with 'Victorian' style produces, against all expectations, not a distinctively homogeneous and predictable diction and form, but more variation than would be expected in a teleological framework. (1999: 164)

From a contemporary Critical Genre Analysis perspective, this phenomenon could be described as genre hybridity. For Bhatia (2017), CGA is a means of demystifying professional practice through the analysis of interdiscursivity as well as intertextuality. Both interdiscursivity and intertextuality are present in the four manuals, and especially in the three later ones. References to genre conventions appear repeatedly in the four texts, i.e. in the preface and dedication subgenres. References and cross-references to various medical authorities are common in three of the four sources (Forster 1829,[16] Anonymous 1832, Black 1840), as opposed to Carlisle (1817), perhaps precisely because it is the earliest manual and the one that helped plant the seeds of the new hybrid genre. These references create text-internal as well

---

15. Here understood to be equivalent to the term "genre", as in Fairclough (1992: 148) and Bhatia (2017).

16. "Sir Anthony Carlysle [*sic*], in his very sensible book on Old Age" (Forster 1829: 34).

as text-external links. They also create discursive networks across time and space and contribute to interdiscursivity.

Finally, when it comes to "adopting creative and innovative aspects of professional communication" (Bhatia 2017: 12), the one-shilling, 39-page anonymous booklet from 1832 certainly does that by offering an almost picaresque mix of medically acceptable information, sound common sense, humour and even libellous tones, in its sparkling final pages on the demise of a quack (Anonymous 1832: 33). The text also demystifies the established medical profession: "[w]ith these instructions, you need never pay a single fee to a Physician, and you may live to a good old age" (Anonymous 1832: 34). The author is also quick to distinguish his own text from quack practices. He places it in a comfortable middle-of-the-range niche, which probably also corresponded to the author's educational, economic and ultimately social position in that period. Anonymous (1832) thus is the only one of the four self-help manuals belonging to a real hybrid genre, written by a supposedly decently educated non-specialist for the non-specialist (yet not a quack, one must agree).[17] The genre would definitely flourish in the late nineteenth and, especially, in the twentieth centuries.

Maci et al.'s (2015) "genre network" is yet another way of defining the four manuals and their sometimes traditional, sometimes hybrid and sometimes new forms of science popularisation. Indeed, the analysis raised the question of who was an expert and could thus produce this genre at a moment in history when science popularisation was still under way (cf. Kramer et al. 2003). Throughout modern history, professional medical knowledge prevails within the monolithic community of practice of doctors, while others less or non-qualified "health operators" practised among the less affluent, with varying degrees of education, knowledge, expertise and success. In this small selection of self-help manuals, a move emerges precisely towards those "newer forms of popularization, by means of which medical science can be disseminated to laymen as well to practitioners not belonging to the same professional sector" (Maci et al. 2015: 10–11).

The social and scientific changes of the nineteenth century were mirrored in the way the four manuals addressed not only diseases and the medical profession, but their audiences and society in general. Carlisle (1817), for example, stands out for his positive stance to old age, writing about the "[v]alue of old people for society" (Ibid.: 3–4), yet perceiving that the "[v]alue of life decreases as age advances" (Ibid.: 94). He was barely 49 at the time of his book's publication and, since he placed the beginning of senility at around 60, he was probably disinterestedly concerned about it. He also put forward the enlightened consideration that

---

17. For a depiction of real quacks, Porter (1989) covers the history of quacks in England up until 1850.

the propagation of general knowledge threatened those doctors who kept their art obscure and mysterious (Ibid.: 21–24), thus advocating a "democratization of knowledge […] where special sites, long training, and esoteric terms seemed necessarily linked to expertise, exclusivity, and hierarchies" (Worboys 2011: 110). Carlisle shows a humane and caring attitude towards the weak and those in pain and saw the aged as needing the same diet as "children, breeding and nursing women" (Carlisle 1817: 25), recognising them as a vulnerable social group. He wrote extensively about a doctor's aim to "prolong life and lessen pain" (Ibid.: 101), following the expository essayistic tradition. He was not afraid of expressing his opinion: doctors are not to decide which end is better for a patient or whether to dose terminal patients with opium (Ibid.: 102). His considerations on end-of-life practices and palliative care are still topical to this day, and deeply connected to the notions of "agency and choice […] central to structuring medicine, the choice of practitioner, and the use of popular medicine and self-medication" (Waddington 2011: 79). Carlisle was enlightened enough to perceive that the medical profession had to open up to the public, if it were to retain its primacy, but he was also traditional enough to believe that self-medication should be based on established professionals' advice rather than on books showing "technical perplexities" (Carlisle 1817: 1), or even plain quackery. He was also a man of his times in his conservative belief that, when faced with the end of a patient's life, doctors have to "wait for the Almighty command – such questions are indeed above human authority" (Carlisle 1817: 101). Forster (1829) in his turn, made many philosophical and religious remarks, reflecting his personal beliefs. He nonetheless also provided insights into his view of what the relationship between medicine and the government should be – "[m]edicine is fit for the learned and worthy of patronage and protection from government" (Forster 1829: iii) – in a period when medical practice was becoming more institutionalised than ever before and turned into socio-medical *dispositifs*[18] (Rose 2001: 51). The most interesting comments, socially, are however those by Black (1840). A renowned doctor of his time, Black was conservative about the medical authorities that came before him, and made ample reference to them in the theoretical part of his book. Here, he voiced his considerations on the different health conditions within different social classes: firstly, both the lower and the higher classes suffer but for different reasons (Black 1840: 111–112); secondly, the middle classes or the "new rich" suffer the most because they indulge in excessive eating thanks to their newly acquired wealth (Ibid.: 109); thirdly, in his opinion, bowel diseases have increased with "the progress of wealth and the altered state of society, as to labour, luxury, and occupation, both mental and bodily" (Ibid.: 239).

---

18. For a definition of *dispositifs*, see Foucault (1980).

These ideas, although presented as established facts, nonetheless are laid out in a critical yet not disdainful or classist manner but, rather, in the patronizing tone of the family doctor. In fact, Black stands out for his profound belief in progress, being convinced that

(6) the complaints […] will become in each case less severe and fatal, from the *increasing intelligence of the people*, and the greater attention that is paid to *cleanliness* and the *economical comforts of life*, conjoined with the *greater diffusion of scientific knowledge* and of an *improved treatment of disease*, among all the members of the profession.   (Black 1840: 240, emphasis added)

What also emerges is his faith in the gatekeeping role of "the profession": Black was the second youngest of the four authors considered, the last to publish his manual and die, in 1867, still he alternates between his conservative attitudes and his trust in – and probably wonder at – the development of medical science and technology. Thus, Black, but also Carlisle, Forster and the Anonymous author, *all* turn out, in varying ways, to be true to their hybrid and fast-evolving century, in which popular medicine, although derived from its learned counterpart, was not a corruption thereof but represented a common understanding of body, health and disease (Waddington 2011: 81).

6.   Conclusions, limits and developments

This chapter has investigated how "old age" was defined and described in nineteenth-century self-help manuals. The answer was that either it was very heterogeneously described, or simply not dealt with at all (see Table 2). As for the other specific research questions about what "ageing well" meant and what conditions were associated with being old, the answers are disparate. They all (but especially Black 1840) insist on gastrointestinal conditions, and this emphasis is also frequent in the rest of the corpus. The suggestions were sometimes specific, as those by Carlisle (1817) and Black (1840) in the medical sense, and those by Forster (1829) in the spiritual sense. However, they are mostly – and especially those by Anonymous (1832) – common-sense "bits and pieces" of advice, best summed up as a restricted diet and moderation in all aspects of life.

One limitation of the present study concerns the selection criteria. The Wellcome Collection, while particularly rich in medical titles, does not include all nineteenth-century self-help health manuals. Secondly, while the search term "old age" retrieved a small but specific corpus, this does not necessarily imply that their *reading* audience consisted of aged people, but most texts were in fact intended for young or adult people who *wanted* to age well. This, of course, does not mean that

the texts do not share a common concern about the ailments of the elderly, but other search terms such as "life" or "longevity" could have been explored. Thirdly, the four sample manuals selected for the detailed analysis are not entirely homogeneous. The heterogeneity, however, characterises this corpus of 18 self-help medical texts. Only two books in the corpus are possibly the closest thing to manuals both *about* and *for* the elderly. One is Carlisle (1817), which was written by a specialist in the ailments of the aged, and is among the earliest in the corpus; for this reason, it probably shows the features of the traditional essay rather than of the manual, not being divided, for example, into topics or chapters or having a contents page. The other is Van Oven (1853), which is similarly written by a medical doctor and aimed at lay readers for self-help, but is neatly divided into chapters and closer in genre to what would be understood as a present-day manual than Carlisle (1817), possibly because it was written later, after the first half of the century. Finally, although the search was carried out on the period between 1800 and 1899, it seems that Categories (1), (2a) and (3) abound in the first half of the century, probably as a result of the Western tradition of all-encompassing and holistic texts that, when dealing with science, also included philosophical reflections. The second half of the century, on the other hand, sees a trend toward specialisation, as evident in Category (2b) or specific self-help manuals. This tendency continued and established itself in the twentieth century, as is well-known, even giving rise to *hyper*-specialisation in medicine.

Future studies could expand the investigation across genres and times, e.g. including specialised/academic treatises or print media articles. The second half of the nineteenth century and the early twentieth century could also be researched in more detail in a quest of finding the definition of "old" – a problem that is as ancient as mankind and as novel as every new phase of the Anthropocene era.

References

Primary sources

Alcott, William Andrus. 1830. *Health without Physic; Or Cordials for Youth, Manhood, and Old Age. Including Maxims, Moral and Facetious; For the Prevention of Disease, and the Attainment of a Long and Vigorous Life*. London: Effingham Wilson.

Anonymous. 1832. *Twenty Minutes' Advice on Diet, Regimen, and Other Matters Connected with Health. With an Essay on Indigestion, and Rules, by the Observance of which, Persons may Live to a Good Old Age, without Consulting any Medical Man. A Non-Medical Treatise*. London: W. Kidd.

Ballin, Ada S. 1893?. *Health & Beauty in Dress from Infancy to Old Age*. London: Flack.

Bernard, Thomas. 1813. *Comforts of Old Age*. Privately printed for distribution.

Black, James. 1840. *A Manual on the Bowels and the Treatment of their Principal Disorders from Infancy to Old Age*. London: Longman.

Browning, John Dudley. 1884. *Our eyes: and How to Preserve Them from Infancy to Old Age with Special Information about Spectacles*. London: Chatto & Windus.

Canton, Alfred. 1851. *The Teeth and Their Preservation: In Infancy and Manhood to Old Age*. London: Hippolyte Bailliere.

Carlisle, Anthony. 1817. *An Essay on the Disorders of Old Age, and on the Means for Prolonging Human Life*. London: Longman.

Forster, Thomas. 1829. *Medicina Simplex, or, Practical Rules for the Preservation of Health: Printed at the Request of some Friends, for Popular Use*. Chelmsford: Meggy and Chalk. Edition consulted: Forster, Thomas 1829/1832. *Medicina Simplex, or, Practical Rules for the Preservation of Health: Printed at the Request of some Friends, for Popular Use*. Chelmsford: Haddon and Fenton.

Imrie, William. 1835. *The Parents' Dental Guide; A Treatise on the Diseases of the Teeth and Gums from Infancy to Old Age; With Observations on Amalgams, Cements, etc. and Remarks on the Construction of Artificial Teeth*. London: John Churchill.

Jones, William. 1858. *Health for the Million: Including Hints on the Physical Training of Youth, and the Treatment of Invalids and Old Age, with Observations on Unhealthy Employments / by the author of How to make home happy, The accidents of life, etc., etc.* London: W. Kent.

Medicus [pseudonym]. 1830. *The Oracle of Health and Long Life, or, Plain Rules for the Attainment and Preservation of Sound Health and Vigorous Old Age: With Rational Instructions for Diet, Regimen, &c. and the Treatment of Dyspepsy or Indigestion, Deduced from Personal Experience, and the Best Authors on Dietetics*. London: Sherwood, Gilbert, & Piper.

Murphy, Joseph. 1811. *A Natural History of the Human Teeth, with a Treatise on Their Diseases from Infancy to Old Age, Adapted for General Information: To Which Are Added Observations on the Physiognomy of the Teeth and the Projecting Chin*. London: J. Callow.

Paris, John Ayrton. 1826. *A Treatise on Diet: with a View to Establish, on Practical Grounds, a System of Rules for the Prevention and Cure of the Diseases Incident to a Disordered State of the Digestive Functions*. London: Underwood.

Rodd, Thomas (ed). 1837. *Catalogue of Books for MDCCCXXXVII. Containing Theology and Ecclesiastical History. Metaphysics, Moral Philosophy and Education*, 377. London: Compton and Ritchie.

Sinclair, John. 1807. *The Code of Health and Longevity. Or, a Concise View, of the Principles Calculated for the Preservation of Health, and the Attainment of Long Life […]*. Edinburgh: A. Constable & co.

Tweed, John. 1820. *Popular Observations on Regimen and Diet: In Which the Nature and Qualities of Our Common Food Are Pointed Out and Explained: Together with Practical Rules and Regulations in regard to Health, Adapted to Various Situations and Circumstances from Infancy to Old Age*. Chelmsford: Meggy and Chalk.

Van Oven, Bernard. 1853. *On the Decline of Life in Health and Disease, Being an Attempt to Investigate the Causes of Longevity: And the Best Means of Attaining a Healthful Old Age*. London: John Churchill.

West, Francis. 1827. *A Familiar Treatise on the Human Eye: Containing Practical Rules That Will Enable All to Judge What Spectacles Are Best Calculated to Preserve Their Eyes to Extreme Old Age. Illustrated with Three Correct Diagrams of the Human Eye*. London: W. Ackrill.

Whitmarsh, William Michael. 1870. *The Finger-post of Life, or, How to Live: Indicating the Most Direct Way to Attain a Hearty Old Age*. London: Kimpton.

Secondary sources

Bhatia, Vijay K. 2017. *Critical Genre Analysis: Investigating Interdiscursive Performance in Professional Practice*. London and New York: Routledge.

Bynum, William F. 1990. *Science and the Practice of Medicine in the Nineteenth Century*. Cambridge: Cambridge University Press.

Cole, R. J. 1952. "Sir Anthony Carlisle, F.R.S. (1768–1840)." *Annals of Science* 8 (3): 255–270. https://doi.org/10.1080/00033795200200202

Fairclough, Norman. 1992. *Discourse and Social Change*. Cambridge: Polity Press.

Foucault, Michel. [1963] 2003. *The Birth of the Clinic. An Archaeology of Medical Perception*. Milton Park, Abingdon: Routledge, Taylor & Francis Group.

Foucault, Michel. 1980. "The Confession of the Flesh." In *Power/Knowledge Selected Interviews and Other Writings 1972–1977*, ed. by Colin Gordon, 194–228. New York: Pantheon books.

Görlach, Manfred. 1999. *English in Nineteenth-Century England: An Introduction*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511627828

Gotti, Maurizio, Stefania Maci, and Michele Sala (eds). 2015. *Insights into Medical Communication*. Bern: Peter Lang. https://doi.org/10.3726/978-3-0351-0844-6

Johnson, Samuel. 1755. *Dictionary of the English Language*. London: W. Strahan.

Jones, Colin, and Roy Porter (eds). [1994] 2001. *Reassessing Foucault. Power, Medicine and the Body*. London and New York: Routledge.

Kramer, Cheryce, Trea Martyn, and Michael Newton (eds). 2003. *Science as Polite Culture*. Volume 1 of *Literature and Science, 1660–1834*. Part I, ed. by Judith Hawley. London: Pickering and Chatto.

Kytö, Merja, Mats Rydén, and Erik Smitterberg. 2006. *Nineteenth-Century English: Stability and Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486944

Lane, Joan. 2001. *A Social History of Medicine. Health, Healing and Disease in England, 1750–1950*. London and New York: Routledge.

Lane, Joan. 2002. "'The Doctor Scolds Me': The Diaries and Correspondence of Patients in Eighteenth Century England." In *Patients and Practitioners: Lay Perceptions of Medicine in Pre-industrial Society*, ed. by Roy Porter, 205–248. Cambridge: Cambridge University Press.

Lawrence, Christopher. 1994. *Medicine in the Making of Modern Britain, 1700–1920*. London: Routledge. https://doi.org/10.4324/9780203305409

Maci, Stefania, Michele Sala, and Maurizio Gotti. 2015. "Introduction." In *Insights into Medical Communication*, ed. by Maurizio Gotti, Stefania Maci, and Michele Sala, 9–24. Bern: Peter Lang.

ODNB = *Oxford Dictionary of National Biography*. 2004. Oxford: Oxford University Press. http://www.oxforddnb.com/

Ottaway, Susannah R. 2004. *The Decline of Life. Old Age in Eighteenth-Century England*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511495977

Porter, Roy. [1985] 2002. "Laymen, Doctors and Medical Knowledge in the Eighteenth Century: The Evidence of *The Gentleman's Magazine*." In *Patients and Practitioners: Lay Perceptions of Medicine in Pre-Industrial Society*, ed. by Roy Porter, 283–314. Cambridge: Cambridge University Press.

Porter, Roy. [1987] 1995. *Disease, Medicine and Society in England, 1550–1860*. 2nd edition. Cambridge: Cambridge University Press.

Porter, Roy. 1989. *Health for Sale: Quackery in England, 1660–1850*. Manchester: Manchester University Press.

Rose, Nikolas. [1994] 2001. "Medicine, History and the Present." In *Reassessing Foucault. Power, Medicine and the Body*, ed. by Colin Jones, and Roy Porter, 48–72. London and New York: Routledge.

Ruston, Sharon. 2005. "Vegetarianism and Vitality in the Work of Thomas Forster, William Lawrence and P. B. Shelley." *Keats-Shelley Journal* 54: 113–132.

Smith, Ginnie. 2002. "Prescribing the Rules of Health: Self-Help and Advice in the Late Eighteenth Century." In *Patients and Practitioners: Lay Perceptions of Medicine in Pre-Industrial Society*, ed. by Roy Porter, 249–282. Cambridge: Cambridge University Press.

Taavitsainen, Irma, Turo Hiltunen, Anu Lehto, Ville Marttila, Päivi Pahta, Maura Ratia, Carla Suhr, and Jukka Tyrkkö. 2014. "Late Modern English Medical Texts 1700–1800: A Corpus for Analyzing Eighteenth-Century Medical English." *ICAME Journal* 38: 137–153. https://doi.org/10.2478/icame-2014-0007

Taavitsainen, Irma, and Turo Hiltunen. 2019. *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.221

Troyansky, David G. 2005. "The Eighteenth Century." In *A History of Old Age*, ed. by Pat Thane, 175–210. London: Thames & Hudson.

Waddington, Keir. 2011. *An Introduction to the Social History of Medicine. Europe since 1500*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.

Worboys, Michael. 2011. "Practice and the Science of Medicine in the Nineteenth Century." *Isis. A Journal of the History of Science Society* 102: 109–115. https://doi.org/10.1086/658660

Yallop, Helen. 2013. *Age and Identity in Eighteenth-Century England*. London; Brookfield, VT: Pickering & Chatto.

The popularization of learned medicine in late seventeenth-century England

Accommodating translation strategies and textual aspects

Giulia Rovelli
University of Insubria

Following from the popularizing work of Nicholas Culpeper (1616–1654), who published an unauthorized translation of the Royal College of Physicians' *Pharmacopoeia* in 1649, the late seventeenth century witnessed the publication of an unprecedented number of vernacular medical texts, many of which were translations. As the sources of most of these texts had been published in Latin by some of the most important medical authors of the time with the intent of circulating them only among the European medical elite, their translation into English had a profound ideological significance, because it rendered them virtually accessible to all who could read. This chapter analyzes five such vernacularizations in order to gain an insight into the specific translation methods and procedures that were adopted by translators to accommodate the specialized subject and language of medicine to an audience of non-specialists.

Keywords: medical vernacularization, medical popularization, translation studies, historical discourse analysis, historical pragmatics

1. Background

The second half of the seventeenth century has long been described as a particularly significant moment in the history of medical popularization. Indeed, following the ideological and political events of the 1640s and 1650s, censorship collapsed, thus resulting in an explosion of the vernacular medical publishing market (Furdell 2002: 59, Fissell 2007: 113, see also Fissell 2011). Opened by Nicholas Culpeper's unlicensed translation of the *Pharmacopoeia Londinensis* (1649), which overtly criticized the Royal College of Physicians' authority and contested their monopoly over medical matters (Sanderson 1999: 33–34), the late seventeenth century witnessed

the publication of an unprecedented number of vernacular medical books (Furdell 2002: 38). While many of these works were original English productions, a considerable percentage (between 10 and 17 percent) turned out to be translations of learned continental works written originally in Latin (Fissell 2007: 113–114 and Rovelli 2018), which at the time still was the international lingua franca of scholarship (Barber 1976: 43), and thus initially intended to be circulated only among the European medical elite. The translation of such significant texts as Johann Vesling's *Syntagma Anatomicum*[1] or Oswald Croll's *Basilica Chymica*[2] aimed at rendering learned medicine accessible to an audience of increasingly confident groups of lay people and practitioners, who were literate, although not university educated, and who started to demand that "edifying knowledge, including medical knowledge, be disseminated to them" (Furdell 2002: 36, see also Wear 2000: 4–5). As the majority of the population had no access to learned physicians, who were often "viewed as expensive and uncharitable" (Wear 2000: 21), and since the household still represented the main arena for medical treatment (Leong and Pennell 2007: 134), vernacularizations helped to "lift the veil of linguistic secrecy from medical arcana", which was achieved through the use of Latin (Furdell 2002: 59). The choice of the vernacular therefore rendered the specialized knowledge of medicine accessible to a wider reading public, thus contributing to what was defined by Sanderson as "the movement towards the democratization of medical knowledge" (1999: 5).

Out of the huge number of English medical texts published from 1649 to 1699, sixty-five can be safely described as first edition translations of medical texts originally written in Latin (Rovelli 2018). The majority of these sources were published in the sixteenth and seventeenth centuries, by some of the most significant figures in the history of medicine, be they starched supporters of traditional medicine like Jean Riolan (Mani 1968: 143), or great innovators like Reigner de Graaf (Porter 1985: 120) and Thomas Sydenham (Anstey 2011: 458–459). Although most of these texts were explicitly aimed at a particular type of reader, be it lay or specialized, about a third of them mentioned both types of audience, which was not particularly unusual – as stated by Wear, "distinctions between lay and medical readerships were blurred and both groups might read works which were ostensibly for the other" (2000: 40–41). However, since the sources of these texts, being printed in Latin, were easily available to all regular physicians, a non-specialized, lay audience, was

---

1.   Johann Vesling's *Syntagma Anatomicum* (1641) came to be "the most widely used anatomical text in Europe during the second half of the seventeenth and first half of the eighteenth centuries" (Ghosh 2014: 1125–1126). It was translated into English in 1653 by Nicholas Culpeper.

2.   Oswald Croll's *Basilica Chymica* (1635) was considered to be "the standard scientific work of iatrochemistry" (Schröder 2008: 471). Its anonymous English translation dates back to 1670.

the most frequently cited one. Notwithstanding this, some texts specifically targeted a more specialized readership, ranging from physicians and students of medicine to other types of practitioners, which included not only surgeons and apothecaries but also midwives. While most of these texts very generically cited common good as the reason why they were translated, some offered more explicit and practical aims, including the improvement of learning, the disclosure of secrets, and also the different ways in which these types of information could have been useful. Although such translations generally covered all of the traditional medical genres, the most frequent were those that may have had a more direct practical application, like RECIPE COLLECTIONS, *receptaria* in Latin, which collected a series of medical recipes both for the prevention of health and the cure of diseases, and TREATISES ON SPECIFIC TOPICS, which frequently dealt with some of the most topical diseases of the time (fevers, women's diseases, etc.), and as such represented useful tools for the maintenance and restoration of health.

2. Aims, corpus and methods

As English medical writing was still in its infancy and the target audience was only minimally composed of specialists, this chapter intends to offer an insight into how translators manipulated their source texts in order to popularize them, by rendering them more accessible for and comprehensible to their heterogeneous target readership.

It therefore provides a close-reading analysis of five randomly selected 12,000-word samples from three RECIPE COLLECTIONS (cf. Rovelli 2019) and two TREATISES ON SPECIFIC DISEASES, by comparing them to their Latin source texts (see Table 1).

Table 1. The analyzed Latin texts and English translations

Genre	Latin source text	English target text
RECIPE COLLECTIONS	Morel: *Formulae Remediorum* (1650)	Anon.: *The Expert Doctor's Dispensatory* (1657a)
	Croll: *Basilica Chymica* (1635) & Hartmann: *Praxis Chymiatrica* (1659)	Anon.: *Basilica Chymica* (1670)
	Sydenham: *Processus Integri* (1692)	Anon.: *The Compleat Method of Curing Almost All Diseases* (1694)
TREATISES ON SPECIFIC DISEASES	Bauderon: *Praxis* (1620)	Anon.: *The Expert Phisician* (1657b)
	Harvey: *De Febribus* (1672)	Anon.: *A Theoretical and Chiefly Practical Treatise of Fevors* (1674)

The texts were randomly selected from a larger corpus (Rovelli 2018), in order to analyze how the two most popular genres of medical writing were translated in three different decades (the 1650s, 1670s and 1690s),[3] as this might provide an insight into the diachronic evolution of these text types, but also of the popularizing strategies which the translators chose to adopt. Availability of their Latin source texts and the total number of pages were also taken into consideration so as to make data more easily comparable.

Following historical pragmatic (Jucker and Taavitsainen 2010) and historical discourse analytic methodologies (Claridge 2017), the study investigates the discursive and linguistic strategies that translators adopted to accommodate to an audience who was literate, but not university educated. All non-literal translation actions encountered in the comparative analysis of target and source texts were therefore manually collected and classified following a slightly adapted version (see Table 2) of the models devised by Vinay and Darbelnet (1958/1995) and Newmark (1988), as used by Alonso-Almeida and Sánchez (2016). As shown in Table 2, this categorization distinguishes all non-literal translation strategies into 12 different translation procedures, which cover both textual and cultural features and involve the processes of elimination (reduction strategies), addition (extension strategies), and adaptation (focalization and substitution strategies) of source text material (Alonso-Almeida and Sánchez 2016: 46). Frequency figures were then calculated for all translation procedures and normalized to 10,000 words so as to make them comparable.

Table 2. Adapted version of Alonso-Almeida and Sánchez's (2016) model for the analysis of translation strategies and procedures

Translation strategies	Translation procedures
Reduction	Omission
	Condensation
	Implicitation
Extension	Addition
	Explicitation
	Amplification
Focalization	Modulation
	Compensation
Substitution	Partial creation
	(Partial) adaptation
	Equivalence
Non-translation	Foreignization

---

3. Since no translation came out in the 1690s for the TREATISES ON SPECIFIC DISEASES, only two texts from this genre were analyzed.

In order to understand their role in the popularization of learned medical knowledge, the translation procedures thus identified were then mapped onto the popularizing techniques described by Gotti (2003), Calsamiglia and van Dijk (2004), Garzone (2006) and López Orellana (2012). These include the use of fewer technical terms, denomination (or designation), definition, reformulation (or paraphrase), periphrasis, analogy (or association), generalization, exemplification, explication, and simplification.

Finally, the chapter also attempts to trace any possible diatypic and diachronic changes in the use of translation procedures and popularizing devices, which may provide an insight into the early evolution of some of the most frequently translated medical genres of the time.

3. Results and discussion

3.1 Macro-textual and discursive elements

Although all five texts are overtly and recognizably presented as translations, something which may also be adduced by their being anonymous, they were also assimilated into the vernacular tradition of popular medical writing in various ways. With the exception of Anon. (1670) and Anon. (1694), which generally closely follow their source texts, title pages accommodate to the new target audience which the translations intended to reach. Anon. (1657a) is the most significant in this sense, as it introduces a variety of elements whose function was to embed the text into the vernacular tradition of popular medicine, as testified by the references to the "book of secrets" (see the chapters in Leong and Rankin 2011) and "popular errors" (Gentilcore 2004) genres (Examples 1 and 2), as well as appeals to accessibility and practical usefulness (Example 3):

(1) "The Apothecaries Shop, and Chyrurgions Closet open'd"
(Anon. 1657a: title page)

(2) "all safe and honest practices are maintained, and dangerous mistakes discovered"
(Anon. 1657a: title page)

(3) "an easie and usefull method for Practice"
(Anon. 1657a: title page)

Anon. (1657a) also contains a cross-reference to the by then very popular work of Nicholas Culpeper, whose scathing critique of the Royal College of Physicians and its monopoly over medical matters had made him the champion of popularization. Indeed, the English title of Anon. (1657a), *The Expert Doctors Dispensatory*, undoubtedly recalls Culpeper's first and most significant publication, namely the

London Dispensatory, the unauthorized translation of the College's *Pharmacopoeia*.[4] The description of the contents, which reads "what out of subtility for their own profits they have indeavoured to reserve to themselves, now at last impartially divulged and made common" (Anon. 1657a: title page), is also in line with Culpeper's democratic ideology of free access to medical notions, as it openly denounces the exclusionary practices of learned physicians, who, according to Culpeper, retain the use of Latin to enslave common people for profit (cf. Cook 1986: 121, Sanderson 1999: 127). Finally, not unlike most vernacular works of the time, the contents of the book are described in much detail, thus further inscribing the text into the English tradition:

(4)   Containing, First the Latine names of all Simples and Compounds English'd. Secondly, The Vertues, Qualities, Properties, Quantities, and uses of all Simples and Compounds. Thirdly, The way of prescribing remedies; together with the Forms and Rules for the making of all manner of medicines, daily used by our English Physicians, Chyrurgians, and Apothecaries. Fourthly, The Nature, Qualities, and Symptomes of all diseases. Fifthly, Cautions for the applying all both internal and external medicines.   (Anon. 1657a: title page)

Anon. (1657b), which is the translation of the first treatise of Bauderon's *Praxis* (1620), also presents quite a few accommodating elements. While the title, *The Expert Physician*, does in no way correspond to that of the source text, the subtitle ("Learnedly treating of all Agues and Feavers, Whether Simple or Compound", Anon. 1657b: title page) clearly follows the Latin one ("De Febribus essentialibus tam simplicitus quam compositis", Bauderon 1620: title page) and, at the same time, renders it more explicit, as Latin 'febribus' is translated into English as "agues and feavers" (Anon. 1657b: title page). Similarly to most vernacular texts of the time, Anon. (1657b) also enumerates the contents of the book in greater detail, thus inscribing the texts into the English tradition:

(5)   A Feverish Heat. The differences of Feavers. A Diary Feaver. A Burning Feaver. A continual Putrid. A continual Tertian. A continual Quotidian. A continual Quartan. An intermitting Quartan. Feavers annexed to Quartans. A Semitertian Feaver. An Hectick Feaver. Confused Erratick Feavers. Malignant pestilent Feavers, &c.   (Anon. 1657b: title page)

---

4.   Although Nicholas Culpeper's original translation of the *Pharmacopoeia Londinensis* (1649) was titled *A Physicall Directory, or A Translation of the London Dispensatory*, the text came to be referred to simply as the "London Dispensatory", the only part of the original title which survived in later editions.

The text also describes its aim as the "general good of this Nation" (Anon. 1657b: title page), thus inserting it into the democratic tradition which acknowledged the need for medical notions, as practical and useful, to be made accessible to all. The practical aspect is also emphasized in the title page, as the instructions contained in the book are described as easy to "put in practice" (Anon. 1657b: title page), thus also implicitly mentioning another very common topos of vernacular medical publishing, namely accessibility.

Adaptation is also present in the prefatory material of the texts, which generally replace the Latin apparatus of dedications and letters with original English material, thus addressing the new target reader and specifically English realities and problems. Counting on Nicholas Culpeper's brand name, which assured readers of the worth of any medical book and was therefore used by publishers to boost the sales of many vernacular works which had very little to do with Culpeper himself (McCarl 1996: 230), Anon. (1657a) also adds "Nicholas Culpeppers Approbation, Or Rather his Wish after his perusal of that Famous, *Morellus* his Dispensatory", in which Nicholas Culpeper himself supposedly (he had died three years before) approves of this "most useful, compendious and exact Dispensatory" (Anon. 1657a:A3), wishing he had translated it. Anon. (1657a) also introduces a letter penned by the bookseller Nathaniel Brook to "all ingenious Practitioners in Physick", a phrase which probably does not include regular physicians, as these were university-trained and thus well-learned in classical languages, and frequently disapproved of such accomplishments. Indeed, as "popular writers were continually challenged for 'prostituting' the secrets of the sciences by publishing translations of Latin works originally meant for academic audiences" (Eamon 2011: 27), Brook seemingly criticizes and distances himself from the much too common custom of having medical texts and authors "surreptitiously, adventurously, and dangerously" translated into English (Anon. 1657a:B). However, while this may be read as an apologetic attitude intended to curry favor with critics, it may also have functioned as an advertising technique to assure vernacular readers that the book was made for them (Fissell 2007: 423).

"The Epistle to the Reader" added in Anon. (1657b) is slightly more explicit as far as popularization is concerned, as it overtly mentions the unfairness of having such a useful and learned work "cloistered up in the French and Latine tongue" (Anon. 1657b:A4), thus inscribing the translation in the popularizing tradition started by Nicholas Culpeper. The accessibility claim, also typical in the vernacularizations of the time, is quite prominent in this translation as well, which is described as "methodical, facil, and perspicuous enough to benefit the meanest capacity, yet satisfie the highest" (Anon. 1657b:A5), which Fissell suggests reading as a rhetorical and advertising strategy, more than a realistic characteristic of the

text, "signaling that not much knowledge nor deep literacy skills were required to use the book" (2007: 111). Notwithstanding this, accessibility does seem to be an important aspect of the translation, as it also adds a very short but useful explanation of how to interpret weights and quantities, i.e. "A Graine is the Quantity of a Barley Corn. A Scruple is twenty Barley Cornes. Three Scruples contain a Dram. Eight Drams contain an Ounce" (Anon. 1657b:A8), which was certainly meant to accommodate to the less specialized but nonetheless interested literate lay audience described by Wear (2000).

Although less conspicuously constructed as an accommodation to its readers, its main purpose being to present a defense of alchemy by proving its usefulness in medicine, Anon. (1670) as well has a new letter to the reader which, albeit less explicitly, still represents an attempt to inscribe the text into the vernacular medical tradition. Indeed, the translator, not unlike his contemporaries, criticizes learned physicians for having censored such an important part of medicine and mentions the topos of accessibility, as the work is described as a "gate for any ingenious intellect, with great facility to enter into Alchimy" (Anon. 1670:a2^r–a3). Moreover, the typographical conventions and translation choices adopted in the text are praised for having "illucidated, enlarged, and rendered [it] more plain, and better to be understood" (*ibid.*). Furthermore, the translator also added a "Postscript to the Reader" in which, for his "better understanding of the Ensuing Work" (Anon. 1670:a3^r), he mentions other works cited in the text, and describes the structure of the treatise.

Accommodation is more limited in Anon. (1674), but still present. Indeed, while this text maintains the Latin letter to the reader as is, it also adds a new letter, this time penned by the bookseller William Thackeray, in which, following the vernacular medical tradition, the aim of the translation is identified in "publick service" (Anon. 1674: the Bookseller to the Reader).

The modifications to the source text of Anon. (1694) are, however, quite conspicuous. While its source (Sydenham 1692) did not include any paratextual and prefatory material, Anon. (1694) has "The Translators Preface", which comments on the text and its usefulness, and inscribes it into the vernacular medical tradition. Similarly to Anon. (1657a), this translation also presents an apology of sorts for the decision of having rendered yet another text in English. Indeed, contrary to the general belief that held vernacular books responsible for the proliferation of quacks and charlatans (Porter 1992: 7, Eamon 2011: 27), it is the physicians' preference for Latin which is here condemned as the reason why patients turn to such figures in the first place:

(6) *All translations of Medicinal Books are by many judged to be not only useless but pernicious; and such as procure 'em to be published in the Vulgar Languages, are accused of no less a Crime, than of doing all they can to furnish Madmen with Weapons to murder themselves, and to expose the Lives of Men to the Mercy of Fool and Knaves. I acknowledge indeed, that the World, and perhaps this Nation more than any other part of it, is exceedingly pestered with Quacks. […] And I verily believe, that there is nothing that hath done greater disservice to the Honour and Interest of Physicians, than their over-carefulness to conceal the Knowledge of the Art which they profess from the Public. […] For Men will be apt to suspect, and think too that they have reason to do so, that there must be a great deal of Mischief, or very little true Worth in that which is hid with so much care, and that they are either afraid or ashamed to expose to light. […] Their Want of Knowledge doth also expose them to the Impostures of Empirics, who, like all other* Juglers, *love to play in the Dark. Thus we see that the Ignorant are the Quacks best Customers, and who is there that hath read the Works of but one* Learned Physician, *that would not tremble to put his Life into the Hands of a* Mountebank.   (Anon. 1694:A4–A5)

By adducing knowledge as a solution to, not the cause of, people's reliance on quacks and mountebanks, the translator condemns physicians' "over-carefulness to conceal" medical notions, which instead of protecting people, keeps them in ignorance and "expose[s] them to the Impostures of Empirics", thus aligning himself to Nicholas Culpeper's democratic ideology. "The Translators Preface" also describes the book as easy "to be read, and understood too, in Hours, by any intelligent Person" (Anon. 1694:A3ʳ), thus also mentioning the topos of accessibility, a practice which was very common at the time, both as an accommodating technique and as an advertising strategy (Fissell 2007: 111).

To sum up, although all the texts under investigation are explicitly presented as translations of some of the most important Latin medical books of the time, the translators also strove to embed them in the vernacular medical tradition, thus offering lay readers a much more familiar reference frame, which they could relate to more easily.

3.2   Translation procedures and popularization

From the point of view of the translation methods, that is, the general way of translating, which may result in a free, literal, communicative, or philological translation (Marco 2009: 70), all five texts may be described as literal translations. Indeed, as also proved by the very low percentages of non-literal translation actions in the samples under scrutiny (from 1% in Anon. 1670 and Anon. 1674, to 6% of

Anon. 1657b),[5] they all closely follow both the structure and content of their STs.[6] However, in line with Burke's (2007: 26) assessment of post-medieval translation practice, translating sense-for-sense, rather than word-for-word, turned out to be the preferred orientation, as evidenced by the relative low frequency of instances of foreignization, the translation strategy whereby a SL element is maintained as is, or partially adapted to the phonological or morphological structures of the TL in the TT (with an average of 16 instances per 10,000 words).

As shown in Figures 1 and 2, amplification, which consists in the use of a larger number of morphemes in the TT as compared to the ST (Alonso-Almeida and Sánchez 2016: 47), is the overall most frequently exploited translation procedure, especially in Anon. (1657b; with a frequency of 252 per 10,000 words), Anon. (1694; 214 per 10,000 words) and Anon. (1657a; 103 per 10,000 words). Such a procedure has three main uses in the texts. Firstly, and especially in RECIPE COLLECTIONS, it is employed to avoid all the abbreviations and symbols which might constitute an obstacle for the literate but non-specialized readership that the translations aimed at reaching, as shown in Examples (7) and (8):

(7) ST: ℞ Coralli <u>rub. & fragm. v. lap.</u> pretiosorum super Porphirium <u>sub.</u> tritorum, <u>singul.</u> Э ɪ[7]   (Bauderon 1620: 48)

TT: <u>Take</u> of <u>red</u> Corral, <u>and</u> the <u>fragments</u> of the <u>five</u> precious <u>Stones finely powdered each a scruple</u>   (Anon. 1657b: 80)

(8) ST: ℞ Aq. Ceras. nigr. Lact. Alexiter. aa.Ʒ iij.   (Sydenham 1692: 3)

TT: <u>Take Black-cherry water, Milk-water, of each three ounces</u>   (Anon. 1694: 1)

Secondly, amplification is used to paraphrase particularly difficult Latin terms or expressions, which are thus rendered more easily comprehensible through the use of periphrases (Examples 9 and 10), or etymological explanations (Example 11):

(9) ST: insipida   (Morel 1650: 425)

TT: almost quite without tast   (Anon. 1657a: 247)

---

5. Percentages were calculated by dividing all instances of non-literal translation strategies (manually classified as described in Section 2) by the total number of words in the samples under scrutiny and, therefore, provide only an estimate of the actual proportion of non-literal translation actions.

6. For the sake of brevity, the following conventions are used throughout this section: ST (source text), SL (source language), SC (source culture), TT (target text), TL (target language) and TC (target culture).

7. Emphasis added in all Examples.
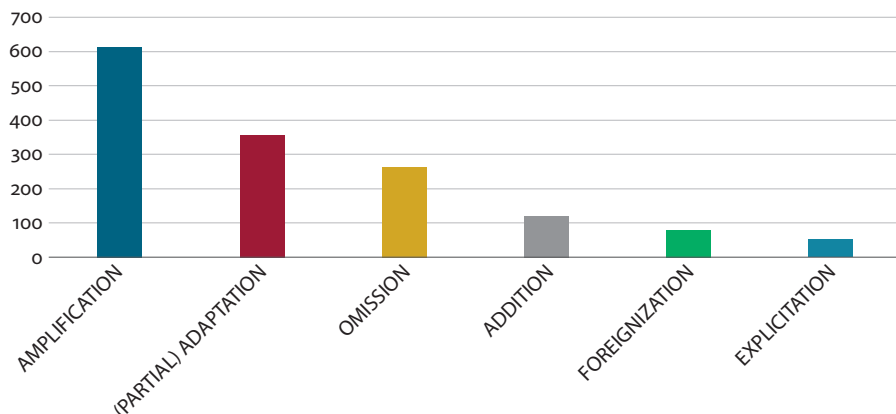
Figure 1. Normalized frequency (per 10,000 words) of the major translation procedures used in the analyzed texts
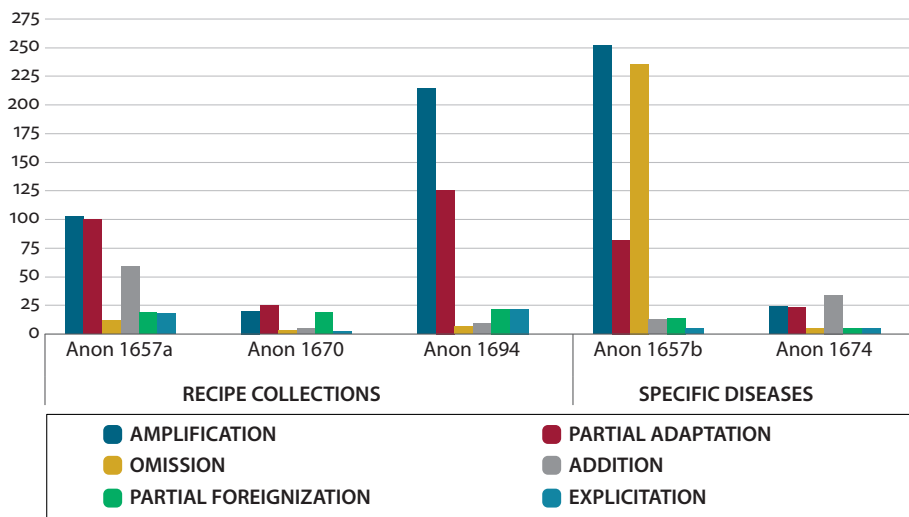
Figure 2. Normalized frequency (per 10,000 words) of the major translation procedures used in each text under scrutiny

(10) ST: in non recutitis   (Sydenham 1692: 42)
      TT: if the Head, or Nut of the Yard be still covered with the foreskin
                                                               (Anon. 1694: 64)

(11) ST: Hepatica   (Bauderon 1620: 8)
      TT: from the Liver Hepatica   (Anon. 1657b: 10)

Finally, amplification is exploited to gloss difficult Latinate terminology with easier synonyms, a strategy which is reminiscent of bilingual dictionaries and glossaries. Its purpose, as shown in Examples (12) through (14), was twofold, as, on the one hand, it rendered the text easier to comprehend for a non-specialized readership, and on the other, it also served to teach the specialized language of medicine to lay readers:

(12) ST: radices scorzonere   (Morel 1650: 512)
      TT: roots of Scorzonera, <u>or Spanish Vipersgrasse</u>   (Anon. 1657a: 431)

(13) ST: intumescentiae   (Harvey 1672: 54)
      TT: intumescence <u>or swelling</u>   (Anon. 1674: 71)

(14) ST: Eclegma   (Sydenham 1692: 4)
      TT: Eclegma <u>(or licking medicine)</u>   (Anon. 1694: 3)

As amplification renders the text more intelligible by providing more comprehensible substitutes for obscure terminology and jargon, be it through periphrases, paraphrases, or explications (cf. Gotti 2003, Calsamiglia and van Dijk 2004, Garzone 2006), its impact on the popularization of specialized medical matters was certainly of the utmost significance. The very high frequency of this procedure, which almost doubles the second most frequent one, is therefore not surprising, especially since such glosses have had a long tradition in English (McConchie 2019: 69).

    The second most frequent translation procedure in the texts is, perhaps unsurprisingly, (partial) adaptation, whereby cultural-specific elements of the SC in the ST are replaced by cultural-specific elements of the TC in the TT (Alonso-Almeida and Sánchez 2016: 47), thus resulting in a preference for native terminology and common language. This procedure is especially frequent in Anon. (1694; with a frequency of 126 per 10,000 words), Anon. (1657a; 100 per 10,000 words) and Anon. (1657b; 82 per 10,000 words), specifically with Latin ingredient names, but also with those of diseases, medicines, procedures and anatomical parts. Although these terms existed in a slightly adapted English form, in the analyzed texts they were replaced by easier equivalents of Germanic origin, as shown in Examples (15) through (19):

(15) ST: anethi   (Morel 1650: 519)
      TT: dill[8]   (Anon. 1657a: 436)

---

8. 'Anet' also existed in English, its first occurrence is recorded by the OED around 1265.

(16) ST: intestinis (Bauderon 1620: 82)
     TT: guts[9] (Anon. 1657b: 108)

(17) ST: Scrophulas (Croll 1635: 379)
     TT: Kings-evil[10] (Anon. 1670: 164)

(18) ST: phlebotomia (Harvey 1672: 52)
     TT: bleeding[11] (Anon. 1674: 68)

(19) ST: enema (Sydenham 1692: 9)
     TT: clyster[12] (Anon. 1694: 10)

As is evident from all the Examples above, this strategy, which ultimately results in the use of fewer technical terms (Gotti 2003: 397), had a huge popularizing import, as words from the general language and of a vernacular origin (López Orellana 2012: 87) were preferred to more obscure Latinate technical expressions.

The third most frequent translation procedure made use of in the texts under scrutiny is omission, that is, the exclusion of a portion of the ST from the TT (Alonso-Almeida and Sánchez 2016: 47). Although to some extent exploited in all texts, especially as far as quotation and citations are concerned, its overall very high frequency is mostly due to Anon. (1657b; 236 per 10,000 words), in which entire stretches of text are often omitted. Nonetheless, the abridgement of long texts, which was customary at the time (Burke 2007: 31), had a great popularizing potential, as it served to simplify learned treatises and, as a consequence, rendered them more accessible for a less sophisticated readership (López Orellana 2012: 87).

The opposite of omission, addition, is the fourth most frequent translation procedure in the corpus, with a particularly high frequency in Anon. (1657a; 59 per 10,000 words) and Anon. (1674; 34 per 10,000 words), where such added material was anticipated in the prefatory text: "I have obtained the favor from the Author […] to insert the addition of very remarkable and important observations and cures" (Anon. 1674: the Bookseller to the Reader), and "these additions where you see there two Commaes (") at the head of the line, were not in the Latine; wherefore seeing it defective most […] where it should have been most large, I added them" (Anon. 1657a: 300). All other translations, however, following the common practice of the time (cf. Burke 2007: 31–32), simply added material without warning the reader. As "modern texts were not infrequently considered capable of improvement by their translators" and modifying a text was by no means

---

9. 'Intestines' also existed in English, the OED records its first attestation in 1598.

10. 'Scrofula' also existed in English, the OED records its first attestation before 1400.

11. 'Phlebotomy' also existed in English, the OED records its first attestation before 1400.

12. 'Enema' also existed in English, although it was first attested only in 1681 (cf. OED).

infrequent (ibid.: 30), additions were probably intended to ameliorate the text, by providing readers with other remedies that the translator thought useful for curing a specific ailment (Example 20), further ingredients that could be used to prepare medical concoctions (Example 21), concrete examples of the success that such remedies have had (Example 22), and also comparisons with more familiar entities (Example 23), which could help readers better understand the text:

(20) R. *Mucilag. sem. cydonior.* ℥iij.
    Suc. plantag. ℥j.
    Mell. rosar. ʒvj.
    Sal. prunel. Əij.
    m. f. Collut.
    quo saepe colluat os.   (Anon. 1674: 109)

(21) Quinces, Pippins, Oranges without the peels, Borage roots
                                                       (Anon. 1657a: 436)

(22) ⸿ I have seen this very successfully and speedily done by a Midwife, only first fomenting the Dug with Vinegar, and afterwards applying a Plaister of Diachilon simple, which was suffered to lye on two or three days; this prevented any hardnesse that otherwise might have happened   (Anon. 1657a: 437)

(23) Tumours like Gums   (Anon. 1670: 20)

Additions, however, are sometimes also used to better structure the TT, by introducing marginal notes (Example [24]) and what might be referred to as section headings (Example [25]), whose function was to, first of all, organize the text and, secondly, guide the reader through it:

(24) The moyst Feaver   (Anon. 1657b: 152)

(25) *The kind.* […] *The Qualities.* […] *The Correction.* […] *The Dose.* […] *The manner of using it.*   (Anon. 1657a: 350)

Finally, additions are also used to provide readers with linguistic notes of a term's equivalent in other languages (especially Greek and Latin), as shown in Example (26) below, thus helping the reader develop a subject-specific register:

(26) Cordials : *by the Latines Cordialia* : *by the Greeks καρδιακα*
                                                   (Anon. 1657a: 430)

The popularizing significance of addition was, therefore, very prominent, as it could be used to enhance readers' comprehension of the text, by giving it a clearer structure which might be easier to follow, by supplying explicit linguistic explanations which could help readers in developing a subject-specific language, but also by linking new and pre-existing knowledge through examples and analogies

which might render specialized medical notions more accessible for an audience of non-specialists (Gotti 2003, Calsamiglia and van Dijk 2004, Garzone 2006).

Although more characteristic of the medieval translating method (Burke 2007: 26), foreignization is also somewhat used in the texts, especially when ingredient and disease names are mentioned, as shown in Examples (27) and (28):

(27) ST: Arsenicum   (Harvey 1672: 67)
      TT: Arsenick[13]   (Anon. 1674: 89)

(28) ST: Erysipelas   (Croll 1635: 384)
      TT: Erysipela's[14]   (Anon. 1670: 166)

While this procedure may have been used to develop readers' specialized language, it could also be related to a lack of knowledge on the translator's part, or to custom, as some Latinate terms might have been more frequent or also more fashionable at the time. The hypothesis that the preference for a vernacular or Latinate equivalent was sometimes dictated by fashion or custom seems to be evidenced by the fact that in the corpus under investigation some terms, like 'Carduus Benedictus'[15] or 'Plantane/Plantain',[16] are always retained in their Latinate form, while others, like 'Mugwort'[17] and 'Wormwood',[18] are always to be found in their adapted form. However, as some terms, like Latin 'aristolochia' and 'mercurius', are variously retained in their Latinate form in some texts, namely 'aristolochy' and 'mercury', and rendered with a vernacular equivalent, specifically 'birthwort' and 'quick-silver' in others, the translators' own preferences also probably played an important role in the choice of translation strategy. Notwithstanding this, some terms, like Latin 'diarrhea' in Anon. (1694), are also translated within the same text following the two different strategies, thus alternatively rendered as 'diarrhea' and 'looseness'. This might, on the one hand, be read as a popularizing technique, in that it provided readers with multiple synonyms of the same term, thus facilitating understanding. On the other hand, since it was in no way indicated that the two terms referred to the same entity, it could be considered a shortcoming of the translation, which might have hindered, rather than fostered, readers' understanding of the text.

---

13. In English also referred to as 'Water-Pepper' (cf. OED).

14. In English also referred to as 'St. Anthony's Fire' or 'The Rose' (cf. OED).

15. In English also referred to as 'Blessed Thistle' (cf. OED).

16. In English also referred to as 'Ribwort' (cf. OED).

17. 'Artemisia' also existed in English; the OED records its first attestation in Anglo-Saxon times.

18. 'Absynth' also existed in English; the OED records its first attestation around 1429.

The last translation procedure which is used in quite a significant number of cases, especially in Anon. (1694; with a frequency of 22 per 10,000 words) and Anon. (1657a; 18 per 10,000 words), is explicitation, that is, the explicit rendering in the TT of implicit information in the ST (Alonso-Almeida and Sánchez 2016: 47). Since in the majority of cases this procedure was used to render the ST less synthetic and more accurate, as shown in Examples (29) and (30), it probably mainly served to ensure readers' access to and comprehension of the text:

(29)   ST: Calida   (Morel 1650: 512)
         TT: The Hot Cordials   (Anon. 1657a: 430)

(30)   ST: pectoralibus   (Sydenham 1692: 56)
         TT: Pectoral Medicines   (Anon. 1694: 94)

The remaining procedures, namely condensation, implicitation, compensation, modulation and partial creation, are only sparingly, if ever, used and, as such, do not provide a real insight into the translation practices of the texts under scrutiny.

3.3   Diatypic variation and translator attitude

The frequency of the translation procedures in the three decades in which the translations were published shows no clear diachronic trends. Indeed, although some rising (as in the case of adaptation) or decreasing tendencies (as for omission or addition) may seem to occur, these are actually caused by a particularly high frequency in one single text and, as such, do not represent a real evolution.
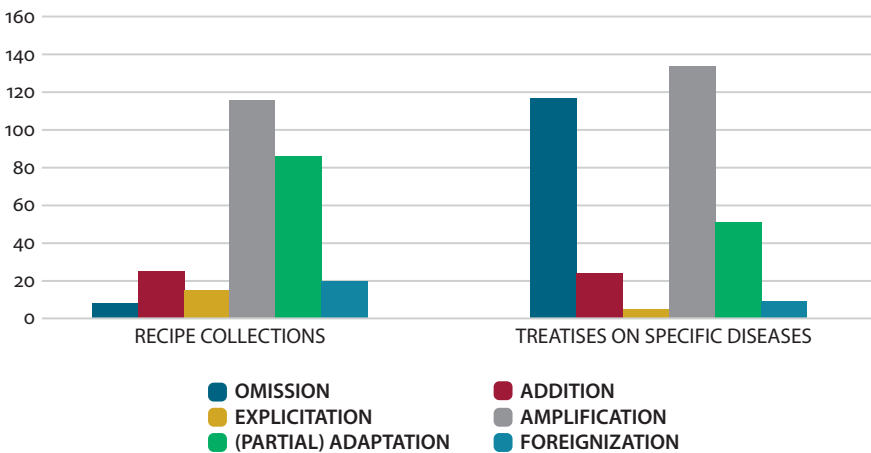
Figure 3. Distribution of the major translation procedures in the two genres of RECIPE COLLECTIONS and TREATISES ON SPECIFIC DISEASES

However, as shown in Figure 3, some possible genre-related characteristics do seem to emerge. Although no real procedure seems to be specific to the TREATISES ON SPECIFIC DISEASES genre, something which may also be conditioned by the limitedness of the data, amplification and (partial) adaptation appear to emerge as more typical of the RECIPE COLLECTIONS. Indeed, their very high frequency in this genre may be attributed to its specific textual characteristics, namely the significant presence of measurements and dosages, whose symbols are generally avoided in translation, and the prominence of abbreviated ingredient names, which tend to be fully written out and accommodated to the vernacular tradition and language.

Despite these common trends, the prevalence of one procedure over the others in general seems to be dictated by the compiler's own attitude and preferences, which might also offer an insight into how the translators themselves perceived their role. For instance, Anon. (1657a) and Anon. (1694) seem to privilege accommodation to their target readers, as evidenced by the consistent use of amplification and adaptation. In contrast, Anon. (1670) appears to adopt a more literal method of translation, with a very low number of non-literal translation actions and the preference for Latinate, as opposed to vernacular, terminology. Finally, Anon. (1657b) seems to embody the prototype of the modern translator, who viewed himself rather as a co-author (Burke 2007: 34), as evidenced by the surprisingly high frequency of both omission and, albeit to a lesser extent, addition, two procedures which heavily modify the ST.

4. Conclusions

The study revealed how literalism, albeit in a sense-for-sense acceptation, generally dominated the early modern culture of translation, thus seeming to confirm Burke's (2007) assessment of post-medieval translation practice. Indeed, the general tendency to follow both the structure and the content of the ST, the relatively low incidence of non-literal translation actions, but also the frequent mention of 'faithfulness' in the prefatory material of the translations under scrutiny, all seem to point to literal translation as the general orientation of late-seventeenth-century medical translators.

However, although due to the limitedness of the data no definitive conclusions can be made, adaptation also seems to have played an important role in the vernacularization of learned medical knowledge, as the translators generally tended to accommodate the texts to an audience of non-specialists, who were certainly literate, but not university-educated. Vernacular terminology was, indeed, generally preferred to the Latinate one, especially in the case of ingredient and disease names, thus producing a text which is easier to comprehend. Difficult terminology was

also very frequently glossed in the text itself, by coupling it with easier synonyms, generally of vernacular origin, or by providing reformulations and periphrases, but also concrete examples and comparisons with more familiar entities, whose aim was to simplify such learned texts and render them more accessible for their target readership.

To conclude, albeit the accommodation strategies adopted in the translation of the five texts under scrutiny are on the whole quite limited in number, they nonetheless played a very significant role in the popularization of learned medicine, as they modified the ST in quite meaningful ways. Indeed, since their purpose was that of creating a version of the texts which could be more comprehensible and acceptable for their target audience, they seem to have played a fundamental role in the movement towards the democratization of learned medicine (Sanderson 1999: 5), as they gave lay readers access to useful knowledge which until that time had been kept from them.

References

Primary sources

Anon. 1657a. *The Expert Doctors Dispensatory. The Whole Art of Physick Restored to Practice*. London: printed for N[athaniel] Brook.

Anon. 1657b. *The Expert Phisician: Learnedly Treating of all Agues and Feavers*. London: printed by R.I. for John Hancock.

Anon. 1670. *Bazilica Chymica, & Praxis Chymiatricae, or, Royal and Practical Chymistry in Three Treatises*. London: printed for John Starkey, and Thomas Passinger.

Anon. 1674. *Theoretical and Chiefly Practical Treatise of Fevors wherein It'S Made Evident That the Modern Practice of Curing Continual Fevors Is Dangerous and Very Unsuccessful*. London: printed for William Thackeray.

Anon. 1694. *The Compleat Method of Curing Almost All Diseases to which is Added an Exact Description of Their Several Symptoms*. London: printed and are to be sold by Randal Taylor.

Bauderon, Brice. 1620. *Praxis in Duos Tractatos Distincta*. Lutetia Parisiorum: Sébastien Cramoisy.

Croll, Oswald. 1635. *Oswaldi Crollii Basilica Chymica, Pluribus Selectis & Secretissimis Propria Manuali Experientia Approbatis Descritionibus, & Usu Remediorum Chymicorum Selectissimourm Aucta*. Geneve: Pierre Chouët.

Culpeper, Nicholas. 1649. *A Physicall Directory or a Translation of the London Dispensatory*. London: Peter Cole.

Hartmann, Johan. 1659. *Iohannis Hartmanni Medicinae Doctoris et Quondam Chymiatriæ in Academia Marpurgensi Professori Celeberrimi, Principumque Hassie Archiatri Praxis Chymiatrica*. Geneve: Samuel Chouët.

Harvey, Gideon. 1672. *De Febribus Tractatus Theoreticus, et Practicus*. London: William Thackeray.

Morel, Pierre. 1650. *Formulæ Remediorum Studio & Opera*. Rouen: Jean Berthelin.

Sydenham, Thomas. 1692. *Processus Integri in Morbis Fere Omnibus Curandis*. London.

Vesling, Johann. 1641. *Syntagma Anatomicum*. Frankfurt: Johannes Beyer.

Secondary sources

Alonso-Almeida, Francisco, and Ángeles Sánchez. 2016. "'If They have not the French': Translation Choices in *The Happy Deliverie of Women* (1612)." *The Translator* 22 (1): 40–58. https://doi.org/10.1080/13556509.2015.1032535

Anstey, Peter. 2011. "The Creation of the English Hippocrates." *Medical History* 55: 457–478. https://doi.org/10.1017/S0025727300004944

Barber, Charles. 1976. *Early Modern English*. London: André Deutsch.

Burke, Peter. 2007. "Cultures of Translation in Early Modern Europe." In *Cultural Translation in Early Modern Europe*, ed. by Peter Burke, and Ronnie Po-Chia Hsia, 7–38. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511497193.002

Calsamiglia, Helena, and Teun A. van Dijk. 2004. "Popularization Discourse and Knowledge about the Genome." *Discourse & Society* 15 (4): 369–389. https://doi.org/10.1177/0957926504043705

Claridge, Claudia. 2017. "Discourse-Based Approaches." In *English Historical Linguistics. Approaches and Perspectives*, ed. by Laurel Brinton, 185–217. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316286562.008

Cook, Harold John. 1986. *The Decline of the Old Medical Regime in Stuart London*. Ithaca/London: Cornell University Press. https://doi.org/10.7591/9781501733840

Eamon, William. 2011. "How to Read a Book of Secrets?" In *Secrets and Knowledge in Medicine and Science, 1500–1800*, ed. by Elaine Leong and Alisha Rankin. Farnham, Surrey: Ashgate.

Fissell, Mary. 2007. "The Marketplace of Print." In *Medicine and the Market in England and its Colonies, c. 1450–c. 1850*, ed. by Mark S. R. Jenner, and Patrick Wallis, 108–132. Basingstoke and New York: Palgrave Macmillan. https://doi.org/10.1057/9780230591462_6

Fissell, Mary. 2011. "Popular Medical Writing." In *The Oxford History of Popular Print Culture: Volume One: Cheap Print in Britain and Ireland to 1660*, ed. by Joad Raymond, 418–431. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199287048.003.0030

Furdell, Elizabeth L. 2002. *Publishing and Medicine in Early Modern England*. Rochester, N.Y.: University of Rochester Press.

Garzone, Giuliana. 2006. *Perspectives on ESP and Popularization*. Milan: CUEM.

Gentilcore, David. 2004. "Was There a 'Popular Medicine' in Early Modern England?" *Folklore* 115 (2): 151–166. https://doi.org/10.1080/0015587042000231255

Ghosh, Sanjib K. 2014. "Johann Vesling (1598–1649): Seventeenth Century Anatomist of Padua and His *Syntagma Anatomicum*." *Clinical Anatomy* 27: 1122–1127. https://doi.org/10.1002/ca.22454

Gotti, Maurizio. 2003. *Specialized Discourse: Linguistic Features and Changing Conventions*. Bern: Peter Lang.

Jucker, Andreas H., and Irma Taavitsainen. 2010. *Historical Pragmatics*. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110214284

Leong, Elaine, and Sarah Pennell. 2007. "Recipe Collections and the Currency of Medical Knowledge in the Early Modern 'Medical Marketplace.'" In *Medicine and the Market in England and Its Colonies, c. 1450–c. 1850*, ed. by Mark S. R. Jenner, and Patrick Wallis, 133–152. Basingstoke and New York: Palgrave Macmillan. https://doi.org/10.1057/9780230591462_7

Leong, Elaine, and Alisha Rankin (eds.). 2011. *Secrets and Knowledge in Medicine and Science, 1500–1800*. Farnham, Surrey: Ashgate.

López Orellana, M. Ángeles. 2012. "Popularising Scientific Discourse." *Quaderns de filología. Estudis lingüístics* 17: 83–96.

Mani, Nikolaus. 1968. "Jean Riolan II (1580–1657) and Medical Research." *Bulletin of the History of Medicine* 42 (2): 121–144.

Marco, Josep. 2009. "The Terminology of Translation. Epistemological, conceptual and intercultural problems and their social consequences." In *The Metalanguage of Translation*, ed. by Yves Gambier, and Luc van Doorslaer, 65–80. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/bct.20.05mar

McCarl, Mary R. 1996. "Publishing the Works of Nicholas Culpeper, Astrological Herbalist and Translator of Latin Medical Works in Seventeenth-Century London". *Canadian Bulletin of Medical History 1996* 13(2), 225–276. https://doi.org/10.3138/cbmh.13.2.225

McConchie, Roderick. 2019. *Discovery in Haste: English Medical Dictionaries and Lexicographers 1547 to 1796*. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110639186

Newmark, Peter. 1988. *A Textbook of Translation*. New York and London: Prentice Hall.

OED. 2015. *The Oxford English Dictionary*. Oxford: Oxford University Press.

Porter, Roy. 1985. "Reigner de Graaf." In *Dizionario Biografico della Medicina e delle Scienze Naturali*, ed. by Roy Porter, 120. Parma: Ricci Editore.

Porter, Roy. 1992. "Introduction". In *The Popularisation of Medicine 1650–1850* ed. by Roy Porter, 1–17. London and New York: Routledge. https://doi.org/10.1017/CBO9781139170215.001

Rovelli, Giulia. 2018. "'For the Benefit of those who Understand not the Latine Tongue'. The Vernacularization of Medicine in Late-Seventeenth-Century England." *Expressio* 2: 123–151.

Rovelli, Giulia. 2019. "Translating Latin Medicine in Late-Seventeenth-Century England: Discursive Aspects and Popularizing Strategies." *Status Quaestionis* 17: 166–195.

Sanderson, Jonathan. 1999. Nicholas Culpeper and the Book Trade: Print and the Promotion of Vernacular Medical Knowledge, 1649–65. PhD dissertation. University of Leeds.

Schröder, Gerald. 2008. "Crollius, Oswald." In *Complete Dictionary of Scientific Biography*, ed. by Charles C. Gillispie, 471–472. Detroit, MI.: Charles Scribner's Sons.

Vinay, Jean-Paul, and Jean Darbelnet. 1958/1995. *Comparative Stylistics of French and English: A Methodology for Translation*, translated and edited by Juan Sager and Marie-Jo Hamel. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/btl.11

Wear, Andrew. 2000. *Knowledge and Practice in English Medicine, 1550–1680*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511612763

Index

The original studies in this volume provide new insights into the history
of medical discourse across centuries in both professional and lay texts.
The central themes deal with changes in medical writing in various societal
and cultural contexts in search for best practices in corpus pragmatics
for future work. Some studies apply quantitative methods of corpus
linguistics and Digital Humanities, others adopt a qualitative, discourse-
analytical perspective, focusing on particular texts, authors or medical
topics, or specific functionally-defined discourse forms such as narratives.
Quantitative and qualitative approaches are mutually complementary
and shed light on different aspects of historical medical discourse.
The methodologies aim at establishing validity and reliability for pragmatic
analysis, taking into account relevant contextual factors and insights
from other fields, such as medical and social history, history of ideas,
and science studies.

John Benjamins Publishing Company