Late Modern English Medical Texts 1700–1800: A corpus for analysing eighteenth-century medical English

Irma Taavitsainen,1 Turo Hiltunen,1 Anu Lehto,1 Ville Marttila,1 Päivi Pahta,2 Maura Ratia,1 Carla Suhr3 and Jukka Tyrkkö2
1 Department of Modern Languages, University of Helsinki
2 School of Language, Translation and Literary Studies, University of Tampere
3 School of Languages and Translation Studies, University of Turku

1 Introduction
If we want to study the history of scientific English, the importance of medical texts can hardly be overstated. However, many aspects of the history of this register are still understudied and poorly understood. This is particularly true of the eighteenth century, which has so far received fairly little scholarly attention, much less than the earlier centuries. This is particularly surprising, given that the eighteenth century is an extremely important period in the history of medical science, representing a transfer from the earlier periods towards the more modern approaches.

In this article, we will describe the corpus of Late Modern English Medical Texts (LMEMT), a new resource to facilitate a systematic study of features of medical writing in the eighteenth century. Our intention is to make the corpus available to researchers with its background information in due course.

2 Building corpora for studying the development of medical writing
Research on the language of eighteenth-century medicine has been impeded by the lack of a standard corpus resource that would be sufficiently large to enable the systematic study of a large variety of linguistic features. There are some studies of individual linguistic features of language use,1 but the primary data on which the results are based is usually not made available to other researchers. We believe that by compiling a specialised corpus and making it available to the research community, it is possible to stimulate research in the area. The experi-
ence of the two previous corpora of medical writing compiled by our team, *Middle English Medical Texts* 1375–1500 (MEMT; see Taavitsainen *et al.* 2005) and *Early Modern English Medical Texts* 1500–1700 (EMEMT; see Taavitsainen *et al.* 2010) supports this view: both have inspired a number of linguistic and philological studies on the history of English medical writing up to the year 1700.²

To further research on eighteenth-century medical English, we are currently compiling a corpus of *Late Modern English Medical Texts* (LMEMT). LMEMT will contain c. two million words and reflect an inclusive view of medicine that covers the domain, including both elite and household practices. In the following sections, we shall discuss some issues related to the compilation.

2.1 **Text selection**

The texts in LMEMT have been systematically selected in order to represent the wide variety of medical texts in the period. The texts have been selected in collaboration with medical historians.³ The main source of corpus texts is the online repository *Eighteenth Century Collection Online* (ECCO), which provides access to facsimile images of eighteenth-century printed texts. Through institutional collaboration with the ECCO Text Creation Partnership (TCP) based in Michigan, we have received some of the texts in XML format. In addition, a number of texts have been obtained by agreements with various repositories, and they have been keyed in. Occasionally, we have also taken advantage of data that has been collected for individual research by project members, provided that they fit the general selection criteria. An example of this is midwifery (see Section 3.1).

2.2 **Text categorisation**

In the same way as in MEMT and EMEMT, the texts are divided into discrete categories to facilitate studies of different sub-registers of medical writing. The LMEMT contains six main categories: General treatises and textbooks, Texts on specific diseases, methods, therapeutic substances and midwifery, Recipe collections, Surgical and anatomical texts, Public health and Periodicals. Figure 1 displays the overall plan and the relative size of each category.
Figure 1: LMEMT categories

The principles used for categorising the texts are similar to those used in EMEMT. In other words, the primary categorisation is based on the topic of the text, or the area of medicine it represents. The field of medical writing becomes increasingly complex between the fifteenth and the eighteenth centuries, but topics provide a solid basis for longer diachronic assessments of this language variety. Figure 2 represents schematically the relationship between the categories used in the three corpora (MEMT, EMEMT and LMEMT). A more detailed description of some of these categories is provided in Section 3 below.
It is important to emphasise that texts within a category may vary – and often do so – with regard to other parameters, for example target audience or the author’s educational background. We encourage corpus users to adjust our categorisation to better suit their individual research questions, and provide ample information and links to further sources to do this (see Section 2.4).

2.3 Annotation

LMEMT will be encoded using Extensible Markup Language (XML; Bray et al. 2008) following the principles and practices recommended in the Guidelines for electronic text encoding and interchange developed by the Text Encoding Initiative Consortium (TEI Consortium 2010). Although widely adopted in digital humanities, TEI XML remains relatively rare in historical corpus compilation. By switching from our own proprietary markup, used in MEMT and EMEMT, to XML, we make a considered decision to favour a universally-known and future-proof standard that will allow, among other things, easy conversions into
other formats such as XHTML for dynamic online presentation. The use of XML will also enable links between the corpus texts and metadata and, later on, makes it possible to extend the information content of the corpus with new layers of annotation such as part-of-speech tagging and more finely developed annotation of discursive elements. At present, the markup serves to extend the expressive resources of the corpus text by representing the paratextual aspects of the printed original (Buzzetti 2009). In addition to the usual annotation of text structure (chapters, headings, paragraphs, etc.), the corpus texts will also be annotated with features relevant to research questions on book history, e.g. typographic features, layout and illustrations.

2.4 Background information
One of our main principles is that the corpus findings should be analysed in their sociohistorical context. For this purpose, the corpus will provide historical and sociolinguistic information about the texts and the people involved in their production. Together with the texts, LMEMT will make available a catalogue providing bibliographical information of each text, a description of its subject matter as well as information about the authors (and translators). The catalogue will also cast light on the intended audience of the text and its importance. As in EMEMT, the catalogue entries provide links to images of the original texts in the Eighteenth Century Collections Online (ECCO) allowing the users to view and study features of typography, layout, and the relationship between images and text (subject to ECCO subscription), and also links to further authorial information in the Oxford Dictionary of National Biography (ODNB; also subject to subscription).

3 Some corpus categories in more detail
In this section we shall take a closer look at some of the categories in the corpus, in order to demonstrate the diversity of medical writing in the eighteenth century, but the text categories of the corpus are not limited to those discussed here. Two of the types discussed here – book-length specialised treatises on midwifery (Section 3.1) and surgical treatises (3.2) – have a long history as central categories of medical writing, whereas the others – texts on medical institutions and public welfare (3.3), articles in medical journals (3.4) and medical writings in a newly launched magazine (3.5) – are relative newcomers in the register of medical writing in the eighteenth century.
3.1 Texts on specific topics: Midwifery texts

Texts on specific topics formed by far the largest category in EMEMT, and their importance continues in the eighteenth century. Cutting-edge science was mainly published in monographs, and these texts show how knowledge disseminated in professional circles. At the same time, treatises on specific topics were also published for lay readers.

The category Texts on specific topics in LMEMT includes various subcategories, one of which is texts on midwifery and reproductive medicine. This collection of texts aims at representing the variety of contemporary writings in an area of knowledge characterised by ongoing radical changes. In addition to advances in knowledge, the period witnessed major changes in the praxis and training of midwifery. Until the seventeenth century, assisting childbirth was almost exclusively the job of female midwives, typically trained in apprenticeship. A new type of male practitioner, known as the man-midwife, had emerged in France and the innovation was quickly introduced in Britain. In line with the scientific method of the period, the men-midwives emphasised the importance of observation and experimentation, including dissections and the development of obstetrical tools. In connection with this change, the eighteenth century also saw the beginning of obstetrics as an institutionalised sub-discipline of medicine. The first British school of midwifery was founded in London in 1725 and the first Chair of midwifery in Edinburgh in 1726 (see e.g. Cody 2005). While the new institutional knowledge and training was mainly accessible to men only, women also continued to practice midwifery, relying on the more traditional forms of knowledge. The innovations were received with a degree of contradiction, which also gave rise to a considerable amount of published controversy.

The subcategory comprises fourteen texts. As in most other sub-categories, the texts represent different types of authors and different readerships, and span the century. Most texts are written by medical professionals operating in this field, including male and female practitioners. The target audiences range from practitioners and students to non-professionals. The new approach to reproductive medicine is represented in LMEMT by several texts. Among the most important ones is *A treatise on the theory and practice of midwifery*, first published in 1752, by William Smellie, one of leading eighteenth-century lecturers in obstetrics. Smellie’s textbook, containing case narratives from all classes of pregnant women, was a runaway hit, which helped to advertise men-midwives’ superiority over their female counterparts as vastly experienced, technologically skilled scientists, and observant, empathetic attendants (Cody 2005: 151-152). Another key text representing the new wave, reporting in detail on new discoveries in reproductive anatomy, is *An anatomical description of the human gravid...*
uterus, and its contents by William Hunter, the leading anatomist and surgeon of the period, published posthumously in 1794. Margaret Stephens’ The domestic midwife (1795) is one of the very few eighteenth-century books on midwifery written by a woman who had access to the new training and ran a successful school of midwifery for women. Continuity from the earlier traditions of midwifery is represented by A complete practice of midwifery (1737) by Sarah Stone, a skilled practicing midwife, who had learned the profession as an apprentice to her mother and was critically outspoken about men-midwives. The sample also includes an extract of Nicholas Culpeper’s A directory for midwives [...] newly corrected from many gross errors, printed in 1701. Culpeper’s book was one of the most popular midwifery manuals of the time, appearing in seventeen editions between 1651 and 1777; an earlier edition of the text is included in EMEMT.

3.2 Surgical and anatomical writing
Surgical texts have a long and interesting history within English medicine (see e.g. Wear 2000: Chapter 5), and eighteenth-century surgical writing continued in many ways the traditions set in the early modern period. Surgery and anatomy had been at the forefront of empirical medicine and this development continued in the late modern period when the iconic surgical theatre emerged as a teaching space. The eighteenth century is also strongly associated in the history of medicine with a newfound interest in amputations, made possible by better equipment and better understanding of infections. Major surgical treatises such as those by William Cheselden and John Hunter were used for decades as textbooks in surgical training.

Traditionally the training of barber-surgeons had been taken care of by guilds and companies and was separated from mainline medical education. A controversial milestone in surgical history was reached in 1745 with the separation of surgeons from the Company of Barber-Surgeons and the formation of the new Company of Surgeons. Contrary to what one might expect, the newly-formed Company was unable to maintain the old system of apprenticeship but equally incapable of organising a new formal system of training or licensing surgical practice. The resulting confusion meant that surgical education remained stagnant for the latter half of the eighteenth century. To what extent these developments are reflected in surgical and anatomical texts is one of the interesting questions for corpus-based research.
3.3 Medical institutions and public welfare

Medical institutions and the thought of public health started to develop in the eighteenth century, and consequently the LMEMT corpus will include a category for texts on institutions and public welfare. The goal in the compilation of the category has been to include a wide variety of topics related to public health, e.g. texts on hospitals and dispensaries, works on medical practitioners and medical education as well as texts on medical societies, legislation and specific public health topics such as sanitation. The category further aims at comprising a spectrum of authors and audiences: some of the texts are written in the name of institutions and addressed to the general public and professionals, while others are written both by medical professionals and non-professionals to institutions such as hospitals and the parliament. Specific issues of public health form a third group; it includes texts both for professional and general audiences.

Numerous hospitals and dispensaries were founded by charities in the eighteenth century, and many texts in this category consider the role of the infirmaries and introduce them especially to the subscribers. In London, five general and eleven specialised hospitals were established during the period, so that at the end of the eighteenth century, there were about twenty hospitals in London and almost thirty in the counties (Berridge 1990: 204; Kilpatrick 1990: 254; Porter 1991: 72). In addition, thirteen dispensaries were set up in London between 1770 and 1800, and there were over twenty general dispensaries outside London by 1800 (Berridge 1990: 206; Kilpatrick 1990: 271). These institutions provided medical care especially for the poor, but they also became important for improving medical education. The new hospitals were dependent on the donations from subscribers, and the texts on hospitals in LMEMT praise the benevolence of the nation, as can be seen in example (1) from John Aikin’s *Thoughts on hospitals* (1771):

(1) … Providence has planted in our natures a benevolent principle, which, without waiting for duty to incite, or reason to approve, inclines us by an involuntary emotion to relieve the distresses of our fellow-creatures, and gives us the purest and most sensible pleasure for our reward. …

We have seen, within less than half a century, numerous edifices arise throughout the kingdom, dedicated to the support of the poor under the severe afflictions of disease and want… and all this without any interference of the civil powers, merely by the generous and disinterested zeal of individuals.

(John Aikin, *Thoughts on hospitals* (1771), p. 5–6)
The author notes the generous and disinterested zeal of individuals in the founding of the voluntary hospitals, and refers to the lack of civil powers in the advancement of public health. In addition to the works that discuss the role of hospitals in general, the LMEMT corpus will include texts on specific hospitals – such as the London Hospital and the Bethlem Hospital – that describe the founding and daily life of the infirmaries in more detail.

Public health was further improved by attempting to prevent epidemics. Although there was no systematic public health movement in the era, there were pioneering authors who started to consider the relation of filthy environment and diseases (Porter 1991: 65). In LMEMT, the relation of unsanitary conditions and diseases is explored for instance in George Borthwick’s pamphlet The method of preventing and removing the causes of infectious diseases (1784), which considers how the spread of infections can be prevented by municipal planning. The importance of hygiene was discovered additionally in institutional settings such as prisons, hospitals and the army. Daniel Peter Layard’s Directions to prevent the contagion of the jail-distemper (1772) suggests that typhus can be prevented by appointing physicians to prisons and by administering officials to take care of cleanliness in prisons and court rooms.

Quantification and early statistical methods were important developments in eighteenth-century medical research and they also became “the quantitative measure of public health and welfare” (Rusnock 2002: 1; see also Rusnock 1995). The LMEMT corpus text by William Black entitled An arithmetical and medical analysis of the diseases and mortality of the human species (1789) was influential in medical arithmetic. Black surveyed London bills of mortality and related the calculations with, for example, geography, climate and age. In addition to medical arithmeticians, specialised hospitals brought a large number of similar patients together, promoting the development towards generalised views of diseases instead of diagnoses of single patients (see e.g. Lane 2001). These improved methods together with the institutional changes and the new views on sanitation all aimed at preventing epidemics (Porter 1991: 73). Hence, the eighteenth century saw important developments in the field of institutions and public health, although a centralised public health movement did not take place until the nineteenth century.

### 3.4 Periodicals

The category Periodicals is different from the other categories in that it is defined by its publication format, not the topic of the texts. In other words, texts in this category have been published in scientific journals, but their topics are not limited to just one area of medical research. The scientific article clearly
grew in importance in the eighteenth century, although the book was still the main vehicle for disseminating scientific knowledge. This development is also reflected in the composition of the LMEMT corpus, where this category accounts for over 20 per cent of the total number of words included.

The category contains articles from two journals, the *Philosophical Transactions of the Royal Society* (*Phil. Trans.*) and the *Edinburgh Medical Journal* (EMJ). As such, the category represents both continuity and innovation. The *Phil. Trans.* is the landmark publication in the history of English empiricism, and it had already been established in the seventeenth century (the EMEMT corpus includes a sample of medical articles from this period; see Hiltunen 2010 for details). During the years after its establishment in the 1660s, the Royal Society was devoted to the study of natural philosophy in various forms, and a substantial number of its members were medical practitioners by occupation (Hall 1971: 112; Porter 1989: 272). The EMJ, by contrast, represents another stage in the history of medical periodicals: transactions of strictly medical societies established in the eighteenth century (Lefanu 1938, Kronick 1976). The first issue of the journal came out in 1733, under the title *Medical Essays and Observations.*

Texts in the category Periodicals make up a substantial portion of all articles published in the two journals in the eighteenth century. As for *Phil. Trans.*, our sample includes roughly one third of all medical articles published in the journal. As the total number of articles in the EMJ is much larger, our sample comprises roughly every twentieth article published in the journal.

Corpus-based research on the history of British scientific periodicals was pioneered by Dwight Atkinson, who has investigated the rhetoric of both the *Phil. Trans.* and the EMJ since their inception to the present day (Atkinson 1992, 1999). Using a combination of rhetorical and multi-dimensional analysis (see Biber 1988), Atkinson described the changes in the rhetoric of medical research over a long period of time, including a shift from individual cases of disease to general disease types, the fall of the author-centred discourse, and the gradual replacing of sequential narratives by ‘modular’ reports, which ultimately pave the way to the modern IMRAD (Introduction, Methods, Results and Discussion) macrostructure, now ubiquitous in medical research reports (see e.g. Piqué-Angordans and Posteguillo 2006). Example (2) offers an illustration of the narrative style characteristic of the period. It follows the chronological order of events and emphasises the patient’s feelings and the doctor’s role in treatment, which is contrary to present-day practices of foregrounding the methods and measurable details of diagnosis and treatment.
(2) CHARLES BURNET, about 12 years old, having, on Thursday evening the 9th of May, been engaged with some of his companions in throwing stones, received a blow on his head, which made him fall down, and he remained insensible for a few minutes. He was soon after led home, and appeared so immaterially hurt, as to require no immediate assistance.

On the following morning I was sent for, and found him out of bed, complaining only of some sickness, giddiness, and inclination for vomiting; his senses every way perfect; his sight and hearing very distinct.

(Thomas Brown: “History of a case in which a recovery took place, after a remarkable injury to the brain”. *EMJ*, vol. 18, p. 342 (1794))

The category *Periodicals* in LMEMT contains a more comprehensive text collection than those used in previous studies. As such, it allows us to take a closer look at the literate developments specific to the medical science after the rise of empiricism in Britain, and more importantly, contrast them with developments in other contemporaneous medical texts in the corpus.

3.5 Gentleman’s Magazine (GM)

Magazines provided a new channel for the distribution of medical knowledge to lay people. The influence of the period style of politeness is evident in these texts. Founded in 1731, *Gentleman’s Magazine* (GM) was a digest from contemporary newspaper writings with miscellaneous materials (Porter 1985; Taavit-sainen forthcoming). Medical items in GM increased towards the end of the century with polemical articles on current issues like public health, seamen’s conditions, longevity and new cures. The magazine also contained book reviews and statistical assessments of weather conditions and diseases, as well as bills of mortality. Letters to the Editor show how medical knowledge was appropriated at large and how lay people felt about issues of health. The following letter in example (3) illustrates this type of writing:

(3) Mr. Urban, Oct. 5.

HAVING been a sufferer from *ascarides* above half a century, and having taken every reasonable remedy I could hear of, from *æthrops mineral* when an infant, without effect, I can give very little comfort to *A Medical Sufferer*, p. 804 except that, as he is not an old man, he may, however, arrive at old age notwithstanding the teasing disorder which he labours under.
I am not of the medical profession, as your correspondent will probably have perceived; but I hope he will accept this information, such as it is, from A Fellow-sufferer.

(GM, p. 892 (1792); emphasis original)

4 Conclusion

The diachronic development of conventions in scientific and medical writing is an extremely rich and fascinating object of study. It is also an important area of research: a full understanding of present-day practices is possible only if we understand how they have come about, know about past conventions and can relate them to the changing patterns of thought in scientific thinking (cf. Crombie 1995).

LMEMT covers a broad range of medical texts from cutting-edge scientific treatises to household handbooks, and its main contribution to the corpus linguistic research community is to provide convenient access to a large and representative database containing materials that have previously received little scholarly attention. Together with the medical corpora that have already been released – EMEMT and MEMT – LMEMT will facilitate further diachronic studies of medical English, making it possible to investigate linguistic processes in the development of the special language of medicine.

Finally, LMEMT also gives users tools to probe into the socio-historical and cultural contexts of texts, and by doing so, enables interdisciplinary research at the interface of corpus linguistics, philology, history of science, and book history. The inclusion of a wide range of texts, both learned and popular, makes it possible to address entirely new research questions combining insights from various fields of study.

Notes

43,000-word sample of medical writing from the eighteenth century. For more details, see http://www.helsinki.fi/varieng/CoRD/corpora/ARCHER/updated%20version/introduction.html.

2. MEMT comprises about half a million words and is structured according to the traditions of writing: Specialised texts, Surgical treatises, and Remedies and materia medica. EMEMT contains a two-million word representative sample of the entire field of English medical writings in print between 1500 and 1700. EMEMT includes texts ranging from theoretical treatises rooted in academic traditions to popularised and utilitarian writing. The texts are grouped into six categories that facilitate systematic research into the history of medical writing in its disciplinary context. Both corpora cover medical writing from highly learned texts to popular genres of writing in household literature, almanacs and pamphlets. See CoRD for descriptions of these corpora, and Lehto et al. (2010) and Taavitsainen and Pahta (2013) for an overview of studies published before 2013 based on them.

3. We would like to acknowledge Peter Murray Jones’s contribution and we are grateful for his advice in selecting the corpus texts. We would also like to thank Elaine Leong for sharing her expertise on household medicine with us.

4. Other categories include recipes, health guides, and medical handbooks for household use. Along with treatises on specific topics, the corpus also contains treatises with a more general scope.

5. To obtain the sample, we first compiled a list of medical articles in Phil. Trans., which we used as the sampling frame. For the first four decades, our selection relies on the collection Medical essays and observations relating to the practice of physic and surgery (1745) by Samuel Mihles. We then randomly selected every third article from each decade.

6. Text samples are randomly selected, and the number of texts included for each decade is proportional to the total number of available texts.

7. Such narratives continue the tradition of medical case studies from earlier periods, starting with late medieval texts (see Taavitsainen 2011 on case studies in MEMT and EMEMT).

8. For instance, Atkinson’s Edinburgh Medical Journal corpus (1735–1985) contains twenty texts from the eighteenth century selected from two volumes, whereas LMEMT will contain over 90 texts that cover the entire period up to the year 1800.
References

**Corpora and electronic resources**


ECCO: *Eighteenth-Century Collections Online*. http://quod.lib.umich.edu/e/ecco/


**Secondary sources**


